

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: Apr 1 7, 2006, 21:24:06 ; Search time 4213 Seconds  
(without alignments) 10038.340 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744  
Sequence: 1 ggcgcagatcattgaagttttt.....aaacgacgtatgctcttcg 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_rsc:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hfg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	6	CS130878
2	40.8	5.5	110000	1	BA000004_31
3	40.2	5.4	239687	14	CR932017
4	39.8	5.3	200322	14	AC164371
5	39.4	5.3	2000	6	AX655393
6	39.4	5.3	204839	14	AC144482
7	38.6	5.2	110000	2	AC116984_2
8	38.6	5.2	110000	2	AC116984_3
9	38.6	5.2	163151	5	BX537280
10	38.6	5.2	244259	14	AC123226
11	38.6	5.2	257204	14	AC109987
12	38.2	5.1	116374	5	AC123226
13	38.2	5.1	242081	14	CR933779
14	37.8	5.1	576	15	AF225410
15	37.8	5.1	1554	15	AY032588
16	37.8	5.1	173237	14	AC155465
17	37.8	5.1	181924	14	AC155611
18	37.8	5.1	200412	14	AC155666

c	19	37.8	5.1	279242	14	AC114079	AC114079	Rattus no
	20	37.8	5.1	346308	14	AC128290	AC128290	Rattus no
	21	37.4	5.0	168293	14	AC018397	AC018397	Homo sapi
	22	37	5.0	39961	8	AC000044	AC000044	Homo sapi
	23	37	5.0	43738	8	AC000034	AC000034	Homo sapi
	24	37	5.0	226007	14	AC103426	AC103426	Rattus no
	25	36.6	4.9	2000	6	AX655393	AX655393	Sequence
c	26	36.6	4.9	2286	15	LR078526	LR078526	Lycopersico
	27	36.4	4.9	176422	9	AC158170	AC158170	Mus muscu
	28	36.2	4.9	387	15	AF123309	AF123309	Gastrodia
	29	36.2	4.9	516	15	GBL277784	GBL277784	Gastrodia
	30	36.2	4.9	687	15	GBL277785	GBL277785	Gastrodia
	31	36.2	4.9	687	15	GBL277786	GBL277786	Gastrodia
	32	36.2	4.9	699	15	GBL277783	GBL277783	Gastrodia
	33	36.2	4.9	2306	15	AF934813	AF934813	Gastrodia
c	34	36.2	4.9	41741	8	AL356500	AL356500	Human DNA
	35	36.2	4.9	110000	15	AP008217_237	AP008217_237	Oryza sat
	36	36.2	4.9	149697	15	AC109832	AC109832	Oryza sat
	37	36.2	4.9	180695	14	AC140661	AC140661	Pan trogl
c	38	36	4.8	63331	14	AC084345	AC084345	Homo sapi
c	39	36	4.8	216589	14	AC091741	AC091741	Homo sapi
c	40	35.6	4.8	26557	8	DQ001128	DQ001128	Homo sapi
c	41	35.6	4.8	79438	8	HSDJ547C9	AL109927	Human DNA
c	42	35.6	4.8	134953	14	AC026088	AC026088	Homo sapi
c	43	35.6	4.8	154350	8	AC016948	AC016948	Homo sapi
	44	35.6	4.8	194418	5	BX001055	BX001055	Zebrafish
c	45	35.6	4.8	197196	14	AC112585	AC112585	Rattus no

#### ALIGNMENTS

RESULT 1	CS130878	744 bp	DNA	linear	PAT 02-AUG-2005
LOCUS	CS130878	Sequence 18 from Patent WO200506339.			
DEFINITION	CS130878				
ACCESSION	CS130878.1	GI:71793146			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	Wilting, R.				
TITLE	Polypeptides of Alicyclobacillus sp.				
JOURNAL	Patent: WO 200506339-A 18 21-JUL-2005;				
Novozymes A/S (DK)					
FEATURES					
source	1..744	Location/Qualifiers			
	/organism="Alicyclobacillus sp."				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:61169"				
	1..744				
	/note="CDS"				
	1..123				
	/note="sig_peptide"				
	124..744				
	/note="mat_peptide"				
ORIGIN					
Query Match	100.0%; Score 744; DB 6; Length 744;				
Best Local Similarity	100.0%; Pred. No. 5.7e-220;				
Matches 744; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	GGCGCAATTATGAAAGTTTGGAGATTTGGTACCGTATATGATGCTGTTATTCAG	60		
DB	1	GGCGCAATTATGAAAGTTTGGAGATTTGGTACCGTATATGATGCTGTTATTCAG	60		
QY	61	TGGGGCGCATGACAGATTTCTGCGTTTGGCGGTTCAATTGTGGCATTATATGCTTC	120		
DB	61	TGGGGCGCATGACAGATTTCTGCGTTTGGCGGTTCAATTGTGGCATTATATGCTTC	120		

```
QY 121 GCGAACACGGTGTATATATTCGAGAAAACACCGCGGAAAGCATCAACGGTAAGCGCT 180
    |||||
DB 121 GCGAACACGGTGTATATATTCGAGAAAACACCGCGGAAAGCATCAACGGTAAGCGCT 180
QY 181 ACACTCTCTTGTTGTTAATTCGACGAAATGTTTACAGGTATGCAAAACGAAAGCAAAATCG 240
    |||||
DB 181 ACACTCTCTTGTTGTTAATTCGACGAAATGTTTACAGGTATGCAAAACGAAAGCAAAATCG 240
QY 241 AGTAGCTCTCCGCTCATATAGTCTAGCAACTGCAACATGCCCAATCAAGCTGCT 300
    |||||
DB 241 AGTAGCTCTCCGCTCATATAGTCTAGCAACTGCAACATGCCCAATCAAGCTGCT 300
QY 301 ACGACTCTCTTCTTCAGTGTGAAAGTTACGATATATCCGTTTACACATACGGAAGGTA 360
    |||||
DB 301 ACGACTCTCTTCTTCAGTGTGAAAGTTACGATATATCCGTTTACACATACGGAAGGTA 360
QY 361 GGAAGCTTGGAATTAATTAAGTTAACTCCCTGCGCAAGTTAAGATGTGGGGTACGACGG 420
    |||||
DB 361 GGAAGCTTGGAATTAATTAAGTTAACTCCCTGCGCAAGTTAAGATGTGGGGTACGACGG 420
QY 421 ATAGGTGAACCGCAAAATGTCGTTTGGTTTATCAACATCAACATAGAAATGACGGA 480
    |||||
DB 421 ATAGGTGAACCGCAAAATGTCGTTTGGTTTATCAACATCAACATAGAAATGACGGA 480
QY 481 TCCACTCTCTAGAGGTGCTGATGACATATTCATTGCAAGACTTAAACGGAAACGTT 540
    |||||
DB 481 TCCACTCTCTAGAGGTGCTGATGACATATTCATTGCAAGACTTAAACGGAAACGTT 540
QY 541 TATCAGCCGGATTTCTACGCTGATATATGCAATTAAGAAATTCAGGACATATTCGACC 600
    |||||
DB 541 TATCAGCCGGATTTCTACGCTGATATATGCAATTAAGAAATTCAGGACATATTCGACC 600
QY 601 GACCTCAACCCGTGTGTGTCATGACGACAAATCTGTTTATGATATGCGGATTTATG 660
    |||||
DB 601 GACCTCAACCCGTGTGTGTCATGACGACAAATCTGTTTATGATATGCGGATTTATG 660
QY 661 ACATATGTGACGTCGGGACGATTACTGCTTGCCTTCAGAGGTTCTTCGCGGTCA 720
    |||||
DB 661 ACATATGTGACGTCGGGACGATTACTGCTTGCCTTCAGAGGTTCTTCGCGGTCA 720
QY 721 GATGAACGACGTATGCTCTTCCG 744
    |||||
DB 721 GATGAACGACGTATGCTCTTCCG 744
```

RESULT 2  
BA000004\_31/c  
Sequence split into 42 fragments LOCUS BA000004 Accession BA000004

Fragment Name	Begin	End
BA000004_00	1	110000
BA000004_01	100001	210000
BA000004_02	200001	310000
BA000004_03	300001	410000
BA000004_04	400001	510000
BA000004_05	500001	610000
BA000004_06	600001	710000
BA000004_07	700001	810000
BA000004_08	800001	910000
BA000004_09	900001	1010000
BA000004_10	1000001	1110000
BA000004_11	1100001	1210000
BA000004_12	1200001	1310000
BA000004_13	1300001	1410000
BA000004_14	1400001	1510000
BA000004_15	1500001	1610000
BA000004_16	1600001	1710000
BA000004_17	1700001	1810000
BA000004_18	1800001	1910000
BA000004_19	1900001	2010000
BA000004_20	2000001	2110000
BA000004_21	2100001	2210000
BA000004_22	2200001	2310000

```
BA000004_23 2300001 2410000
BA000004_24 2400001 2510000
BA000004_25 2500001 2610000
BA000004_26 2600001 2710000
BA000004_27 2700001 2810000
BA000004_28 2800001 2910000
BA000004_29 2900001 3010000
BA000004_30 3000001 3110000
BA000004_31 3100001 3210000
BA000004_32 3200001 3310000
BA000004_33 3300001 3410000
BA000004_34 3400001 3510000
BA000004_35 3500001 3610000
BA000004_36 3600001 3710000
BA000004_37 3700001 3810000
BA000004_38 3800001 3910000
BA000004_39 3900001 4010000
BA000004_40 4000001 4110000
BA000004_41 4100001 4202352
Continuation (32 of 42) of BA000004 from base 3100001 (BA000004 Bacillus halodurans C-12)
```

Query Match 5.5%; Score 40.8; DB 1; Length 110000;  
Beet Local Similarity 51.8%; Pred. No. 0.74; Indels 3; Gaps 1;  
Matches 118; Conservative 0; Mismatches 107;

```
QY 426 TGAACCCGAATGTCGCTTGGTTATCAACATCAACATAGAAATGACGATCCAC 485
    |||||
DB 87150 TGACAGGCGAAGAAAGAACTTCTTATTTGATGATGATACATCAAAATGAAAGGACGA 87091
QY 486 TCTATGAGAGTCTGTTATGACATATTCATTGCAAGACTTAAACGGAAACGTTATCA 545
    |||||
DB 87090 CTCTATTCAGGTGATGTCGCTTCTTTAAATTAAGTCCGAGATG---TAGAATATGA 87034
QY 546 GCCGATTTCTACTGTCGATATATGCAATTAACAAATTCAAGGACTATTCGACGACCT 605
    |||||
DB 87033 TTCTGATTCCTGTGTGACCTTACGCAATGAAAGCTGACCTTTTCTTACTTAAGCT 86974
QY 606 CAACCTGTGTGTCATGACGACAAATCTGTTTATGATATGCGGCA 653
    |||||
DB 86973 AAATCTGTGTTTGAACATCAACAGATGAAGGTATTTATATGTTCCGCA 86926
```

RESULT 3  
CR932017  
LOCUS  
DEFINITION  
Danio rerio clone DKEX-33P9, WORKING DRAFT SEQUENCE, 7 unordered  
pieces.  
ACCESSION  
CR932017.4 GI:66392865  
VERSION  
HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULITOP.  
KEYWORDS  
SOURCE  
Danio rerio (zebrafish)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Osteichthyes;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE  
1 (bases 1 to 239687)  
McLaren, S.  
Direct Submission  
Submitted (21-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zf1sh-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On May 22, 2005 this sequence version replaced gi:56797647.  
JOURNAL  
TITLE  
Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: zf1sh-help@sanger.ac.uk  
Project Information  
Center project name: zK33P9  
Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 235298 bases at least Q40

Consensus quality: 235833 bases at least Q30  
Consensus quality: 236380 bases at least Q20  
Insert size: 239087; sum-of-contigs  
Insert size: 217257; 9.8% error; agarose-fp  
Quality coverage: 9.53x in Q20 bases; sum-of-contigs Quality  
coverage: 10.66x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 5067: contig of 5067 bp in length  
5068 5167: gap of 100 bp  
5168 41822: contig of 36655 bp in length  
41823 41922: gap of 100 bp  
41923 131937: contig of 90015 bp in length  
131938 132037: gap of 100 bp  
132038 143502: contig of 11465 bp in length  
143503 208913: gap of 100 bp  
208913 209013: contig of 65311 bp in length  
209013 236615: gap of 100 bp  
236615 236715: gap of 100 bp  
236715 239687: contig of 2972 bp in length.

## FEATURES

## source

1. 239687  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone\_lib="DKEY-33P9"  
/clone\_1ib="Dantiokey"  
1. 5067  
/note="assembly fragment:00012  
fragment chain:1"  
misc\_feature  
5168. 41822  
/note="assembly fragment:00593  
fragment chain:1"  
misc\_feature  
41923. 131937  
/note="assembly fragment:02166  
fragment chain:1"  
misc\_feature  
132038. 143502  
/note="assembly fragment:00090  
fragment chain:1"  
misc\_feature  
143603. 208913  
/note="assembly fragment:01089  
fragment chain:2"  
misc\_feature  
209014. 236615  
/note="assembly fragment:00175  
fragment chain:2"  
misc\_feature  
236716. 239687  
/note="assembly fragment:00050.0"

## ORIGIN

Query Match 5.4%; Score 40.2; DB 14; Length 239687;  
Best Local Similarity 56.4%; Pred. No. 1.2;  
Matches 75; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 508 AATATTCATTGACAACTTAACGGGAACGTTTATCAGCCGAGTCTTACGCTGAGATA 567  
Db 152572 AATATTCAAATTAAGTTCTTACACAGGAATTAATTAACCAAAATTAAGTTTGAATATA 152631  
QY 568 TATGCAATATACAAATTCAGGAGCTATTCGACGACGACCTCAACCGTGTGTGCATGAGC 627  
Db 152632 TTTTCGTTAGAAATTCACAGAAATTTTTCATAGAACTCAACCTTGTCTTATTAATTTAC 152691  
QY 628 ACAAAATCTCGTAT 640  
Db 152692 TGAATTTCTGGCAT 152704

## RESULT 4

## AC164371

## LOCUS

## DEFINITION

AC164371 200322 bp DNA 1linear HTG 01-JUL-2005  
Bos taurus clone CH240-11714, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 19  
unordered pieces.

## ACCESSION

AC164371.2

## VERSION

GI:68299997

## KEYWORDS

HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ENRICHED.

## SOURCE

Bos taurus

## ORGANISM

Bos taurus

Eumylotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Buthera; laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.

## REFERENCE

## AUTHORS

1 (bases 1 to 200322)

Muzny,D.,Marle,M.,Metzger,M.,Lee,A.,Abramson,S.,Adams,C.,Alder,J.,  
Allen,C.,Allen,H.,Alsdorfs,S.,Amin,A.,Anguiano,D.,  
Anyalebechi,V.,Ayagi,A.,Ayodeji,M.,Bacca,B.,Baden,H.,  
Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,  
Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,  
Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderson,E.,  
Cardenas,V.,Carter,K.,Cavazos,I.,Cesear,H.,Center,A.,  
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,  
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Crete,A.,D'Souza,L.,  
Davila,M.L.,Davis,C.,Day-Carroll,L.,De Anda,C.,Dederich,D.,  
Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,  
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duval,B.,Eaves,K.,  
Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,  
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,  
Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garner,T.,Garza,M.,  
Gebregorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,M.,  
Gunaratne,P.,Haaland,W.,Hamil,C.,Hamilton,K.,Hamilton,K.,  
Harvey,Y.,Havlik,P.,Hawes,S.,Henderson,N.,Hernandez,J.,  
Herrandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,  
Hollins,B.,Howells,S.,Hulik,S.,Hume,T.,Idlebird,D.,Jackson,A.,  
Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,  
Karpach,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Koyar,C.,  
Kowals,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,  
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,  
Lorenshewa,L.,Louisege,H.,Lozano,R.J.,Lu,X.,Ma,D.,  
Maheshwari,M.,Mahindartne,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,  
Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,  
Mawhinney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,B.,  
Miosavljivic,A.,Miner,G.,Minja,E.,Montenayor,J.,Moore,S.,  
Morgan,M.,Morris,K.,Morris,S.,Mundasa,M.,Murphy,M.,Nair,L.,  
Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,  
Nwackeleme,O.,Okwomou,G.,Olampunagoun,A.,Pal,S.,Parks,K.,  
Pasternak,S.,Paul,H.,Perez,A.,Perez,L.,Pfankoch,C.,  
Ploppel,F.,Poindexter,A.,Popovic,D.,Primus,R.,Fu,L.-L.,  
Punzo,M.,Quiroz,J.,Rachlin,B.,Reeves,K.,Regier,M.A.,Reigh,R.,  
Reilly,B.,Reilly,M.,Ren,Y.,Reuter,M.,Richards,S.,Riggs,F.,  
Rives,C.,Rodkey,T.,Rojas,A.,Rose,M.,Rose,R.,Rulz,S.J.,  
Sander,M.,Savery,G.,Scherer,S.,Scott,G.,Shatman,S.,Shen,H.,  
Shetty,J.,Shvartsbeyn,A.,Sisson,I.,Sliter,C.D.,Smajic,D.,  
Sneed,A.,Sodergren,B.,Song,X.-Z.,Sorelle,K.,Soza,J.,  
Steinle,M.,Strong,R.,Sutton,A.,Svatek,A.,Tabot,P.,Taylor,C.,  
Taylor,T.,Thomas,N.,Thomas,S.,Tingey,A.,Tregos,Z.,Umani,K.,  
Valas,R.,Vera,V.,Villasana,D.,Walton,L.,Walker,B.,Wang,J.,  
Wang,O.,Wang,S.,Warren,J.,Warren,R.,Wei,X.,White,F.,  
Williams,G.,Willson,R.,Wiczek,R.,Wooden,H.,Worley,K.,  
Wright,D.,Wright,R.,Wu,J.,Yakub,S.,Yen,J.,Yoon,L.,Yoon,V.,  
Yu,F.,Zhang,J.,Zhou,J.,Zhou,X.,Zhao,S.,Zhou,D.,von  
Niederhausern,A.,Weiss,R.,Smith,D.R.,Holt,R.A.,Smith,H.O.,  
Weinstock,G. and Gibbs,R.A.

Direct Submission  
Unpublished  
2 (bases 1 to 200322)  
Worley,K.C.  
Direct Submission  
Submitted (20-JUN-2005) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 200322)

## REFERENCE

AUTHORS  
TITLE  
JOURNAL

## COMMENT

Cow Genome Sequencing Consortium.

Direct Submission

Submitted (01-JUL-2005) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jun 29, 2005 this sequence version replaced gi:67972940.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rxt/). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: FHSA

Center clone name: CH240-11714

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 190643 bases at least Q40

Consensus quality: 193630 bases at least Q30

Consensus quality: 195478 bases at least Q20

Estimated insert size: 198954; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 19 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence.

\* as soon as it is available and the accession number will

\* be preserved.

```
1 5877: contig of 5877 bp in length
* 5878 6322: gap of 445 bp
* 6323 18172: contig of 11850 bp in length
* 18173 18222: gap of 50 bp
* 18223 20526: contig of 2304 bp in length
* 20527 20577: gap of 50 bp
* 20578 35742: contig of 15166 bp in length
* 35743 36222: gap of 480 bp
* 36223 38281: contig of 2059 bp in length
* 38282 38381: gap of unknown length
* 38382 41147: contig of 2766 bp in length
* 41148 41197: gap of 50 bp
* 41198 55437: contig of 14240 bp in length
* 55438 55562: gap of 125 bp
* 55563 86280: contig of 30718 bp in length
* 86281 86330: gap of 50 bp
* 86331 92028: contig of 5698 bp in length
* 92029 92078: gap of 50 bp
* 92079 119970: contig of 27892 bp in length
* 119971 120047: gap of 77 bp
* 120048 123375: contig of 3328 bp in length
* 123376 123665: gap of 290 bp
* 123666 142980: contig of 19315 bp in length
* 142981 143080: gap of unknown length
* 143081 144260: contig of 1180 bp in length
* 144261 145079: gap of 819 bp
* 145080 192359: contig of 47280 bp in length
* 192360 192459: gap of unknown length
* 192460 193549: contig of 1090 bp in length
```

FEATURES  
source

```
* 193550 193649: gap of unknown length
* 193650 195025: contig of 1376 bp in length
* 195026 195125: gap of unknown length
* 195126 196204: contig of 1079 bp in length
* 196205 196304: gap of unknown length
* 196305 197805: contig of 1501 bp in length
* 197806 197905: gap of unknown length
* 197906 200322: contig of 2417 bp in length.
```

location/Qualifiers

1..200322

/organism="Bos taurus"

/mol\_type="genomic DNA"

/db\_xref="taxon:9913"

/clone="CH240-11714"

5878..6322

/estimated\_length=445

18173..18222

/estimated\_length=50

20527..20576

/estimated\_length=50

35743..36222

/estimated\_length=480

38282..38381

/estimated\_length=unknown

41148..41197

/estimated\_length=50

55438..55562

/estimated\_length=125

86281..86330

/estimated\_length=50

92029..92078

/estimated\_length=50

119971..120047

/estimated\_length=77

123376..123665

/estimated\_length=290

142981..143080

/estimated\_length=unknown

144261..145079

/estimated\_length=819

192360..192459

/estimated\_length=unknown

193550..193649

/estimated\_length=unknown

195026..195125

/estimated\_length=unknown

196205..196304

/estimated\_length=unknown

197806..197905

/estimated\_length=unknown

## ORIGIN

Query Match 5.3%; Score 39.8; DB 14; Length 200322;

Best Local Similarity 60.7%; Pred. No. 1.5;

Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Oy 322 AAGTACGATATATCCGCTTTCACACATACGGAAGTAGAGACTTGAATTAGATT 381
Db 13752 AAGTAGACATATGACAGCTCTTCTCCGGAAGTAGAGACTTGAATTAGAGGA 13811
Oy 382 AACTCCCTGACAGCAAGTTAAGAGTGTGGGTACGACGAGATAGTGA 428
Db 13812 AAAACCTCTGTGATGATGATCAAAATGTGGGCTCTCGGGAAGGGA 13858
```

## RESULT 5

AX655393/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

```
AX655393 2000 bp DNA linear PAT 22-MAR-2003
Sequence 5263 from Patent WO03000898.
AX655393.1 GI:29158207
Oryza sativa
```



ORGANISM	Query Match	5.3%	Score 39.4	DB 6	Length 2000
Oryza sativa	Best Local Similarity	9.1%	Pred. 1.7		
Bacteria	Matches	61	Conservative 304	Mismatches 300	Indels 4
Archaea					Gaps 2
Eukaryota					
Viridiplantae					
Streptophyta					
Embryophyta					
Tracheophyta					
Spermatophyta					
Magnoliophyta					
Liliopsida					
Poales					
Poaceae					
Enarthroidae					
Oryzaceae					
Oryza					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
SOURCE					
ORIGIN					
Query Match					
Best Local Similarity					
Matches					
Conservative					
Mismatches					
Indels					
Gaps					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					
Length					
Score					
Pred.					
DB					

DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Gorilla gorilla gorilla clone CH255-154N17, WORKING DRAFT SEQUENCE, 2 ordered pieces.	AC144882	AC144882.2	GI:31880084	HTG: HTGS PHASE2: HTGS DRAFT.	
Gorilla gorilla gorilla (lowland gorilla)					
Eukaryote; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Gorilla.					
REFERENCE					
AUTHORS					
1 (bases 1 to 204839)					
Antomellis,A., Aylele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bonifard,G.G., Brinkley,C., Brooks,S., Cariaga,K., Chu,G., Coleman,B., Coleman,H., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-U., Hu,P., Hurle,B., Idol,J.R., Karlins,E., Khong,P., Latic,P., Lee-Jin,S.-O., Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C., Masketi,B., McDowell,J., Paguirigan,C., Pearson,R., Portnoy,W.B., Prasad,A., Reddik-Dugue,N., Schandler,K., Schuster,M.G., Shah,K., Simon,C., Stanciripop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.					
NISC Comparative Sequencing Initiative					
Unpublished					
2 (bases 1 to 204839)					
Green,E.D.					
REFERENCE					
AUTHORS					
Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717					
Groveomt Circle, Gathersburg, MD 20877, USA					
3 (bases 1 to 204839)					
Green,E.D.					
REFERENCE					
AUTHORS					
Submitted (18-JUN-2003) NIH Intramural Sequencing Center, 8717					
Groveomt Circle, Gathersburg, MD 20877, USA					
On Jun 18, 2003 this sequence version replaced gi:31044297.					
COMMENT					
-----					
Genome Center					
Center: NIH Intramural Sequencing Center					
Center code: NISC					
Web site: <a href="http://www.nisc.nih.gov">http://www.nisc.nih.gov</a>					
Contact: <a href="mailto:nisc.zoo@nih.gov">nisc.zoo@nih.gov</a>					
-----					
Project Information					
Center project name: enf					
Center clone name: 154N17					
-----					
The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.					
-----					
Summary Statistics					
Sequencing vector: plasmid; n/a; 100% of reads					
Chemistry: Dye-terminator Big Dye; 100% of reads					
Assembly program: Phrap; version 0.990319					
Consensus quality: 204590 bases at least Q40					
Consensus quality: 204680 bases at least Q30					
Consensus quality: 204709 bases at least Q20					
Insert size: 192000; agarose-fp					
Insert size: 204739; sum-of-contigs					
Quality coverage: 11.99x in Q20 bases; agarose-fp					
Quality coverage: 11.24x in Q20 bases; sum-of-contigs					
-----					
* NOTE: This is a 'working draft' sequence. It currently					
* consists of 2 contigs. Gaps between the contigs					
* are represented as runs of N. The order of the pieces					
* is believed to be correct as given, however the sizes					
* of the gaps between them are based on estimates that have					
* provided by the submittor.					



The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TrEMBL, Wp: WormPEP, Information on the WormPEP database can be found at [http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhixiong Rao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, Mashu). For further information see [http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) DKEX-228C11 is from a Zebrafish BAC library

## FEATURES

## source

1.163151  
Location/Qualifiers  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEX-228C11"  
/clone\_1ib="DanioKey"

## ORIGIN

Query Match 5.2%; Score 38.6; DB 5; Length 163151;  
Best Local Similarity 46.7%; Pired 30.6; Indels 0; Gaps 0;  
Matches 122; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

5 GATTATGAAAGTTTGGATGATTTGGTACCGATATCATGCGTTATTCAGTGGG 64

Db 139459 GAAAGATGATATTTTGGGCTGATGCTGGCATTTTGTCTTATCCGATATTCATTGGG 139518

65 GCGGAATGAACAGATTCGCGTTTGGCGGTTATGTTGGGCGATTATGTTCTTCGCGCA 124

Db 139519 AAGGTGCTATCCAGATGAGACCTGTCGAGAACGCTGCTCACTTAAATGTTAGTTTTCG 139578

Qy 125 AACAGGTGATATGATTCAGAGAAACACCGCGAAGCATCATCAGTTAGCGCTACAA 184

Db 139579 AAGATTCGTGCTTATCACTGGGTGACAAATACCTGTAACCATTTATATATATATAT 139638

Qy 185 CTCTTGTGTTATTCAGCAATATGTTACAGGTAGTGAAGCAAGCAAACTGAGTA 244

Db 139639 AATATGATATCAATTTTAACTTTTATAGACAAAACAACCTCTGTAACCATGCGATGTC 139698

Qy 245 CGTCTCCGCTCATAGTCTA 265

Db 139699 ATTCCCTCAGCTGMAATGCTA 139719

## RESULT 10

AC123226 244259 bp DNA linear HTG 08-OCT-2002

LOCUS AC123226/c Rattus norvegicus clone CH230-217M18, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\* 8 unordered pieces.

AC123226 GI:23195011

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.

WORDS Rattus norvegicus (Norway rat)

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Schirognathi; Murioidea; Muridae; Murinae; Rattus.

1 (bases 1 to 244259)

Muzny, D., Marie, M., Metzger, M., Lee, S., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alspbrooks, S., Amin, A., Angiano, D.,

Anyalebchi, V., Ayagi, A., Ayodeji, M., Baca, B., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benamed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Burch, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

## TITLE

Unpublished

2 (bases 1 to 244259)

Worley, K.C.

Direct Submission

Submitted (29-MAY-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 244259)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (08-OCT-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Sep 19, 2002 this sequence version replaced gi:21671648.

The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.bgsc.tmc.edu/projects/atlas/>). Each contig described

in the feature table below represents a scaffold in the Atlas

assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated

by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome

shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.bgsc.tmc.edu/>

Contact: bgsc-help@bcm.tmc.edu

----- Project Information



TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
Weinstock, G. and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 257204)  
Worley, K.C.  
Direct Submission  
Submitted (09-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 257204)  
Rat Genome Sequencing Consortium.  
Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Oct 11, 2002 this sequence version replaced gi:21739166.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GRIS  
Center clone name: CH230-230B23  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 231067 bases at least Q40  
Consensus quality: 235027 bases at least Q30  
Consensus quality: 237586 bases at least Q20  
Estimated insert size: 229826; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 13 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
2601: contig of 2601 bp in length  
2602  
2701: gap of unknown length  
2702  
232156: contig of 229455 bp in length  
232157  
232256: gap of unknown length  
232257  
241482: contig of 9226 bp in length  
241483  
241582: gap of unknown length  
241583  
242755: contig of 1173 bp in length  
242756  
242855: gap of unknown length  
242856  
243963: contig of 1108 bp in length  
243964  
244063: gap of unknown length  
244064  
245276: contig of 1213 bp in length  
245277  
245376: gap of unknown length  
245377  
246513: contig of 1137 bp in length  
246514  
246613: gap of unknown length  
246614  
248232: contig of 1619 bp in length  
248233  
248332: gap of unknown length  
248333  
249982: contig of 1650 bp in length  
249983  
250082: gap of unknown length  
250083  
252270: contig of 2188 bp in length

FEATURES  
source  
1..257204  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-230B23"  
1..1057  
/note="wgs\_end\_extension  
clone\_end:5p6"  
2602..2701  
/estimated\_length=unknown  
2702..3930  
/note="wgs\_end\_extension  
clone\_end:5p6"  
3251..3632  
/note="clone\_boundary  
clone\_end:5p6  
site:BCORI  
end\_sequence:RWB0G12TV"  
220094..220968  
/note="clone\_boundary  
clone\_end:T7  
site:BCORI  
end\_sequence:RWB0G12TV"  
230583..232156  
/note="wgs\_end\_extension  
clone\_end:T7"  
232157..232256  
/estimated\_length=unknown  
241483..241582  
/estimated\_length=unknown  
242756..242855  
/estimated\_length=unknown  
243964..244063  
/estimated\_length=unknown  
245277..245376  
/estimated\_length=unknown  
246514..246613  
/estimated\_length=unknown  
248233..248332  
/estimated\_length=unknown  
249983..250082  
/estimated\_length=unknown  
252271..252370  
/estimated\_length=unknown  
253472..253571  
/estimated\_length=unknown  
255720..255819  
/estimated\_length=unknown  
gap  
/estimated\_length=unknown  
ORIGIN  
Query Match 5.2%; Score 38.6; DB 14; Length 257204;  
Best Local Similarity 54.6%; Pred. No. 3.7;  
Matches 77; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 504 TGGCAATTCATTGCGAAGCTTAACGGGACGTTATCAGCCGGAATTCAGTCTGA 563  
DB 28833 TGGCAATTCGGAATTCGAGCGCTTATGTAAGAGGTAGCCATCTGAACCTACCTCTTA 28892  
QY 564 GATATATCAATATGAATTCAGAGGACATTCGACCGACCTCAACCTCGTGTTCAT 623  
DB 28893 CATTTTCCAGAGCGCATTCATACACGAGCCATACCCCAACCCATATGTGTAT 28952  
QY 624 GACGACAATCTCGATTGGA 644  
DB 28953 AGTTTAAAGACCTACATTGA 28973

RESULT 12  
BX571959/c  
LOCUS  
DEFINITION BX571959 116374 bp DNA linear VRT 05-NOV-2004  
zebrafish DNA sequence from clone DKEX-4616 in linkage group 12.  
complete sequence.  
ACCESSION BX571959  
VERSION BX571959.8 GI:55467255  
KEYWORDS HTG.  
SOURCE Dario rerio (zebrafish)  
ORGANISM Dario rerio  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 116374)  
REFERENCE Barker D.  
TITLE Direct Submission  
AUTHORS Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 5, 2004 this sequence version replaced gi:54304257.  
JOURNAL  
COMMENT  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [zf1sh-help@sanger.ac.uk](mailto:zf1sh-help@sanger.ac.uk)  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMBPP; Information  
on the WORMBPP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpe/Clone-derived](http://www.sanger.ac.uk/Projects/C_elegans/wormpe/Clone-derived)  
zebrafish pUC subclones occasionally display inconsistency over the  
length of mononucleotide A/T runs and conserved TA repeats. Where  
this is found the longest good quality representation will be  
submitted.  
Repeat names beginning 'Dr' were identified by the Recon repeat  
discovery system (Zhirong Bao and Sean Rddy, submitted), and those  
beginning 'dir' were identified by Rick Waterman (Stephen Johnson  
lab, WashU). For further information see  
[http://www.sanger.ac.uk/Projects/D\\_rerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml) DKEX-4616  
is from a zebrafish BAC library  
VECTOR: pindigobac-5.  
FEATURES  
source  
1..116374  
/organism="Dario rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/clone="DKEX-4616"  
/clone\_lib="DanioKey"  
ORIGIN  
Query Match 5.1%; Score 38.2; DB 5; Length 116374;  
Best Local Similarity 59.8%; Pred. No. 4.8;  
Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
QY 508 AATTCATTTGCAAGACTTAAACGGAGCGTTATCAGCGGATTTCTACTGCTAGATA 567  
|||||  
DB 47200 AATTCCAATTTAGTTCTTACGAGGAATTTACTTACCAATATATTAGTTTGATATA 47141  
|||||

QY 568 TATGCAATTAACAATTCCAGGACTATTCGACCGACCTCAACCTCG 614  
|||||  
DB 47140 TTTTCTGTAGAAATTCACAGATATTTTCTATAGAACTCAACCTTG 47094  
|||||  
RESULT 13  
CR933779/c  
LOCUS  
DEFINITION CR933779 242081 bp DNA linear HTG 10-JUN-2005  
Dario rerio chromosome 12 clone DKEX-1305, WORKING DRAFT SEQUENCE,  
12 unordered pieces.  
ACCESSION CR933779  
VERSION CR933779.4 GI:67509338  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_FULFILLTOP.  
SOURCE Dario rerio (zebrafish)  
ORGANISM Dario rerio  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 242081)  
REFERENCE McLaren S.  
TITLE Direct Submission  
AUTHORS Submitted (08-JUN-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[zf1sh-help@sanger.ac.uk](mailto:zf1sh-help@sanger.ac.uk) Clone requests:  
[http://www.sanger.ac.uk/Projects/D\\_rerio/fags.shtml#dataeighc](http://www.sanger.ac.uk/Projects/D_rerio/fags.shtml#dataeighc)  
On Jun 10, 2005 this sequence version replaced gi:58330714.  
JOURNAL  
COMMENT  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [zf1sh-help@sanger.ac.uk](mailto:zf1sh-help@sanger.ac.uk)  
----- Project Information  
Center project name: zK1305  
----- Summary Statistics  
Assembly Program: XGAP; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 236648 bases at least Q40  
Consensus quality: 237713 bases at least Q30  
Consensus quality: 238743 bases at least Q20  
Insert size: 240981; sum-of-contigs  
Insert size: 168181; 37.4% error; agarose-fp  
Quality coverage: 6.79x in Q20 bases; sum-of-contigs Quality  
coverage: 9.84x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
3579 3578: contig of 3578 bp in length  
3679 3678: gap of 100 bp  
13686 13686: contig of 10008 bp in length  
13786 13786: gap of 100 bp  
13787 13787: gap of 100 bp  
45507 45507: contig of 31721 bp in length  
45508 45507: gap of 100 bp  
45608 45607: gap of 100 bp  
60434 60434: contig of 14827 bp in length  
60435 60434: gap of 100 bp  
60535 60534: gap of 100 bp  
72346 72346: contig of 11812 bp in length  
72347 72346: gap of 100 bp  
72447 72446: gap of 100 bp  
79048 79048: contig of 6602 bp in length  
79049 79048: gap of 100 bp  
79149 79148: gap of 100 bp  
90278 90278: contig of 11130 bp in length  
90279 90278: gap of 100 bp  
131546 131546: contig of 41168 bp in length  
131547 131546: gap of 100 bp  
131647 131646: gap of 965 bp in length  
141612 141611: contig of 965 bp in length  
141712 141711: gap of 100 bp  
173884 173884: contig of 32173 bp in length  
173885 173984: gap of 100 bp

\* 173985 237730: contig of 63746 bp in length  
\* 237731 237830: gap of 100 bp  
\* 237831 242081: contig of 4251 bp in length.

FEATURES  
source  
1. 242081  
/organism="Danio rerio"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7955"  
/chromosome="12"  
/clone\_id="DKX1-1305"  
/clone\_lib="DantolKey"  
1. .3578  
/note="assembly fragment:01599  
fragment chain:1"  
misc\_feature  
3679. .13686  
/note="assembly fragment:01661  
fragment chain:1"  
misc\_feature  
13787. .45507  
/note="assembly fragment:02271  
fragment chain:1"  
misc\_feature  
45608. .60434  
/note="assembly fragment:02110  
fragment chain:1"  
misc\_feature  
60535. .72346  
/note="assembly fragment:01984  
fragment chain:1"  
misc\_feature  
72447. .79048  
/note="assembly fragment:01793  
fragment chain:1"  
misc\_feature  
79149. .90278  
/note="assembly fragment:01867  
fragment chain:1"  
misc\_feature  
90379. .131546  
/note="assembly fragment:00001  
fragment chain:1"  
misc\_feature  
131647. .141611  
/note="assembly fragment:01720  
fragment chain:2"  
misc\_feature  
141712. .173884  
/note="assembly fragment:00406  
fragment chain:2"  
misc\_feature  
173985. .237730  
/note="assembly fragment:00838.0"  
237831. .242081  
/note="assembly fragment:01615"  
ORIGIN  
Query Match 5.1%; Score 38.2; DB 14; Length 242081;  
Best Local Similarity 59.8%; Pred. No. 4.9; 43; Indels 0; Gaps 0;  
Matches 64; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 508 ATATTCATTGTCAGAACTTAACGGGAAGTTTATCAGCCGATTTCTACTGCTGAGATA 567  
|||||  
Db 220878 ATATTCGAATTTAGTTCTTGAAGAAATTTACTACAGATATTAAGTTTGAATTA 220819  
|||||

QY 568 TATGCAAAATCAATTCAGGACTATTCGACCGACTCAACCTCG 614  
|||||  
Db 220818 TTTTCGTGAATTCACAGATATTTTCATAGAACTCAACCTTGG 220772  
|||||

RESULT 14  
AF225410 576 bp mRNA linear PLN 21-JUL-2005  
DEFINITION Gastridia elata antifungal protein precursor, mRNA, complete cds.  
ACCESSION AF225410  
VERSION AF225410.2 GI:71040863  
KEYWORDS  
SOURCE Gastridia elata  
ORGANISM Gastridia elata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;  
Epilendroidae; lower Epilendroidae; Gastridiaceae; Gastridia.  
REFERENCE  
1 (bases 1 to 576)

AUTHORS Wang, Y., Li, W. and Guo, S.  
TITLE Purification, in vitro activity and cDNA cloning of Gastridia  
antifungal protein (GAFP) from Gastridia (Gastridia elata)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 576)  
AUTHORS Wang, Y., Li, W., Fu, R., Guo, S. and Sun, Y.  
TITLE Direct Submision  
JOURNAL Submitted (14-JAN-2000) Genetics, Institute of Genetics, Datan  
Road, Beijing 100101, China  
REFERENCE 3 (bases 1 to 576)  
AUTHORS Wang, Y., Li, W., Fu, R., Guo, S. and Sun, Y.  
TITLE Direct Submision  
JOURNAL Submitted (20-JUL-2005) Genetics, Institute of Genetics, Datan  
Road, Beijing 100101, China  
REMARK Sequence update by submitter  
COMMENT On Jul 20, 2005 this sequence version replaced gi:12330253.  
FEATURES  
source  
1. .576  
/organism="Gastridia elata"  
/mol\_type="mRNA"  
/db\_xref="taxon:91201"  
/tissue\_type="secondary corn"  
1. .540  
/codon\_start=1  
/product="antifungal protein precursor"  
/protein\_id="AF225410.2"  
/db\_xref="GI:71040864"  
/translation="MAASASTAVILFFAVTTMMSIALPAPASDRINSGLDPTGSL  
AOGGFLIIONDCNVLVDNNAVMAASGNGASGCVLMQNDGNLVYSGRAIVAS  
NTNRONGNYLLIORDRNVLYIDNSNNMIMATHTVGNALITAIPIHNSOTPAASGAO  
NKVNYLISMYRSRIAG"  
1. .84  
1. .837  
/product="antifungal protein"  
ORIGIN  
Query Match 5.1%; Score 37.8; DB 15; Length 576;  
Best Local Similarity 58.4%; Pred. No. 5.2; 47; Indels 0; Gaps 0;  
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 379 GTTAACCTCCCTCAGCAAGTAAAGTGGGTACGAGGATGTAACCGCAAT 438  
|||||  
Db 325 GGTAATCTACTATCTATCTTCAAGAGATCTGAACGTGTCTATATCAATATTTAAAT 384  
|||||

QY 439 GGTGCGTTTGGGTATCAACATCAACCAATAGAAATGACGATCCATCTCTAT 491  
|||||  
Db 385 AATGCGATTTGGGCAACCAACCAACGTTGAAATGCTGAATATCACTGCCAT 437  
|||||

RESULT 15  
AY032588 1554 bp DNA linear PLN 11-JUN-2001  
LOCUS Gastridia elata antifungal protein gene, complete cds.  
DEFINITION AY032588  
ACCESSION AY032588  
VERSION AY032588.1 GI:14346011  
KEYWORDS  
SOURCE Gastridia elata  
ORGANISM Gastridia elata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;  
Epilendroidae; lower Epilendroidae; Gastridiaceae; Gastridia.  
REFERENCE 1 (bases 1 to 1554)  
AUTHORS Sun, Y.-R., Wang, Y.-Q., Li, W.-B., Zhang, L.-M., Niu, H.-Y. and  
Sun, Y.-R.  
TITLE Gastridia elata antifungal protein gene  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1554)  
AUTHORS Sun, Y.-R., Wang, Y.-Q., Li, W.-B., Zhang, L.-M., Niu, H.-Y. and  
Sun, Y.-R.  
TITLE Direct Submission  
JOURNAL Submitted (15-APR-2001) Group 603, Institute of Genetics, Da Tun  
Road, Beijing 100101, People's Republic of China  
FEATURES  
Location/Qualifiers

source 1..1554  
/organism="Gastrodia elata"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:91201"  
misc\_feature 1..581  
/note="contains promoter and 5' UTR"  
<582..>1091  
/product="antifungal protein"  
582..1091  
CDS  
/codon\_start=1  
/product="antifungal protein"  
/protein\_id="AK59994.1"  
/db\_xref="GI:14346012"

translation="MASPASAVIFLPVAALMSLAMPALAAASOLNAGOTLGTGSL  
AOGPDQVIONDNVLVDNRRVWVWAGTNGKAGSCVLRMORDGNLVIYSGSRVTWAS  
NTNRRDNYILLQRDRVVVIYDSSNNAIWAITGVGNAATVIFHSNGTAAAGAAQ  
NKVNEYLRP"

ORIGIN

Query Match 5.1%; Score 37.8; DB 15; Length 1554;  
Best Local Similarity 58.4%; Pred. No. 5.4;  
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 379 GTTAACCTCCCTGCAGCACTTAAGAGTGTGGGTTACGACGGGATAGTGAAACGCAAT 438  
DB 906 GATACTACTATCTGCTCTCTCAGAGAACGTTAAGTTGTCATATACGATCTTCAAT 965  
QY 439 GGTGGTTTGGGTTATCAACATCAACATTAAGAAATGACGGATCCACTCTTAT 491  
DB 966 AATGCAATTGGGCAACCGGCAACGTTGAATGCTGCAATCACTGTAT 1018

Search completed: April 7, 2006, 22:44:12  
Job time : 4222 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 21:06:31 ; Search time 597 Seconds  
(without alignments)  
8447.242 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744  
Sequence: 1 gtgcgaattatgaagtttt.....aaacgacgtatgctcttcg 744

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333246308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	14	ABE45544 Alicyclob
2	744	100.0	744	14	ABE48730 AEB45582 Alicyclob
3	621	83.5	621	14	ABE45582 Alicyclob
4	39.4	5.3	2000	8	ADA71938 Rice gene
5	36.6	4.9	2000	8	ADA71938 Rice gene
6	36.2	4.9	2306	8	ACC47595 Gastridia
7	36.2	4.9	110000	9	ADA13411.1 Continuation (2 of
8	36.2	4.9	110000	12	ADQ59446.1 Continuation (2 of
9	36.2	4.9	110000	14	ADZ13757.1 Aba17217 Human ner
10	35.6	4.8	7636	5	ABE45582 Alicyclob
11	35.2	4.7	580	10	ADK56798 Plant DNA
12	35	4.7	2349	13	ADK56798 Plant DNA
13	35	4.7	42488	4	AAK65772 Bacterial
14	34	4.6	2000	11	ACL37108 Human imm
15	33.8	4.5	729	13	ADK12170 Rice stre
16	33.8	4.5	860	13	ADK12259 Plant ful
17	33.4	4.5	706	10	ADK56319 Plant ful
18	33.4	4.5	1572	3	AAZ53068 Neisseria
19	33.4	4.5	1848	10	ABZ38315 N. gonorr

20	33.4	4.5	1851	3	AAZ53069	AAZ53069 Neisseria
21	33.4	4.5	12077	3	AAAB1734	AAAB1734 N. mening
22	33.4	4.5	110000	3	AAAB1489_6	Continuation (7 of
23	33.4	4.5	34980	3	AAZ1612	AAZ1612 Neisseria
24	33.2	4.5	7480	13	ADP25286	ADP25286 PRO polyp
25	33.2	4.5	7480	13	ADY17856	ADY17856 DNA encod
26	32.6	4.4	2700	5	AAZ14385	AAZ14385 Mouse bon
27	32.4	4.4	813	8	ACA28717	ACA28717 Prokaryot
28	32.4	4.4	18595	8	AAZ3411	AAZ3411 DNA encod
29	32.4	4.4	49999	2	AAZ23899	AAZ23899 Human LOB
30	32.4	4.4	301477	13	ABD33362	ABD33362 Human can
31	32.2	4.3	123526	4	ADJ79962	ADJ79962 Human gii
32	32	4.3	600	2	AAV86185	AAV86185 EST clone
33	32	4.3	26147	4	ABE11324	ABE11324 Drosophila
34	31.8	4.3	503	6	ABQ50691	ABQ50691 Oligonuc
35	31.8	4.3	503	6	ABQ50690	ABQ50690 Oligonuc
36	31.8	4.3	534	6	ABQ55323	ABQ55323 Human ova
37	31.8	4.3	1317	4	AAE61101	AAE61101 P. putida
38	31.8	4.3	83432	14	ABE61124.3	Continuation (4 of
39	31.8	4.3	110000	14	ADZ45062_07	Continuation (8 of
40	31.6	4.2	633	4	AAZ22538	AAZ22538 Human CDN
41	31.6	4.2	704	10	ADK56823	ADK56823 Plant DNA
42	31.6	4.2	1007	10	ADK56823	ADK56823 Plant DNA
43	31.6	4.2	2000	11	ACL35887	ACL35887 Rice stre
44	31.6	4.2	110000	10	ACF67367_48	Continuation (49 o
45	31.6	4.2	110000	10	ACF67367_48	Continuation (49 o

## ALIGNMENTS

RESULT 1	ABE45544	standard; DNA; 744 BP.
ID	ABE45544	standard; DNA; 744 BP.
XX	ABE45544;	
DT	22-SEP-2005	(first entry)
XX	Alicyclobacillus sp. functional polypeptide DNA, SEQ ID NO: 18.	
DB	Alicyclobacillus sp. functional polypeptide DNA, SEQ ID NO: 18.	
XX	Feedstuff; food; detergent; surfactant; pulp; functional polypeptide;	
KW	gene; ds.	
XX	Alicyclobacillus sp.; DSM 15716.	
OS	Alicyclobacillus sp. functional polypeptide	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..744
FT		/*tag= b
FT		/product= "Alicyclobacillus sp. functional polypeptide"
FT		/transl_except= (pos:1..3, aa:met)
FT		/partial
FT		/notes= "No start and stop codons"
FT	sig_peptide	1..123
FT		/*tag= a
FT	mat_peptide	124..744
FT		/*tag= c
FT		/product= "Alicyclobacillus sp. mature functional polypeptide"
XX	US2005147983-A1.	
XX	07-JUL-2005.	
XX	23-FEB-2004; 2004US-00784592.	
XX	06-JAN-2004; 2004DK-00000010.	
XX	04-FEB-2004; 2004DK-00000015.	
XX	(NOVO ) NOVOZYMES AS.	
XX	Wiltling R, Lassen SF, Ostergaard PR;	
XX		

DR WP1; 2005-511773/52.  
 DR P-PSDB; AEB45569.  
 PT New functional polypeptides having function and amino acid sequence  
 PT similar to known specific bacterial enzymes useful in industrial,  
 PT research and household applications e.g. detergents and food.  
 PS Example 3; SEQ ID NO 18; 83bp; English.  
 XX  
 CC The present invention relates to functional polypeptides encoded by  
 CC polynucleotides comprised in the genome of *Allicyclobacillus* sp. DSM  
 CC 15716. The invention is useful in industrial, research and household  
 CC processes such as in composition to prepare feed and food, in detergent  
 CC formulations and for treating lignocellulosic fabric and pulp. The  
 CC present sequence is *Allicyclobacillus* sp. DSM 15716 functional polypeptide  
 CC DNA.  
 XX  
 SQ Sequence 744 BP; 202 A; 159 C; 179 G; 204 T; 0 U; 0 Other;

Query Match 100.0%; Score 744; DB 14; Length 744;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-237;  
 Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGCAATTATGAAAGTTTGGAGTATTTGGTACCGTATATCATGCTGTTTATTCAG 60  
 DB 1 GTGGCAATTATGAAAGTTTGGAGTATTTGGTACCGTATATCATGCTGTTTATTCAG 60  
 QY 61 TGGGGGCGAATGAAAGAAATCTGCGTTTGGCGGTTTCATTGTGGGCAATTATGCTTC 120  
 DB 61 TGGGGGCGAATGAAAGAAATCTGCGTTTGGCGGTTTCATTGTGGGCAATTATGCTTC 120  
 QY 121 GCGAACAGCGTGTATATGATTCGAGAAACACACCGCGGAAAGCATCAACGTAACGCT 180  
 DB 121 GCGAACAGCGTGTATATGATTCGAGAAACACACCGCGGAAAGCATCAACGTAACGCT 180  
 QY 181 ACAACTCTTGGTTGTTAATTCGACGAATGTTCAAGTAGCAAAAGCAAGCAAAATCTCG 240  
 DB 181 ACAACTCTTGGTTGTTAATTCGACGAATGTTCAAGTAGCAAAAGCAAGCAAAATCTCG 240  
 QY 241 AGTAGGTCCTCCCGTCATAGTCTAGCACTCATGCAAGTTCAGCAATGCAATCACTCTCT 300  
 DB 241 AGTAGGTCCTCCCGTCATAGTCTAGCACTCATGCAAGTTCAGCAATGCAATCACTCTCT 300  
 QY 301 ACGACTTCATCTTCACTGCAAGTTCAGATATATCCGTTTCAACATACGGAAGGTA 360  
 DB 301 ACGACTTCATCTTCACTGCAAGTTCAGATATATCCGTTTCAACATACGGAAGGTA 360  
 QY 361 GGAAGCTTGGAATTAAGATTAACTCCGCGACAGCAAGTTAAGAGTGGGGTACGACGGG 420  
 DB 361 GGAAGCTTGGAATTAAGATTAACTCCGCGACAGCAAGTTAAGAGTGGGGTACGACGGG 420  
 QY 421 ATAGGTGAACCGGCAATGTCGTTGGGTTATCAACATCAATGAAGAAATACCGGA 480  
 DB 421 ATAGGTGAACCGGCAATGTCGTTGGGTTATCAACATCAATGAAGAAATACCGGA 480  
 QY 481 TCCACTCTATGAGAGTGTGATGAGCATATTCATTGACAGAACTTAAACGGGAAGCTT 540  
 DB 481 TCCACTCTATGAGAGTGTGATGAGCATATTCATTGACAGAACTTAAACGGGAAGCTT 540  
 QY 541 TATCAGCCGGATTCTAGCTGAGATATATGCAAAATTCAGGAGCTATTCGAGCC 600  
 DB 541 TATCAGCCGGATTCTAGCTGAGATATATGCAAAATTCAGGAGCTATTCGAGCC 600  
 QY 601 GACCTCAACCCCTGGTGTGCATGACGAATAATCTCGATATTTGAAATGCGGATTTTATG 660  
 DB 601 GACCTCAACCCCTGGTGTGCATGACGAATAATCTCGATATTTGAAATGCGGATTTTATG 660  
 QY 661 ACATATGTCAGCTGCGGAGCATTAATCACTTGTGCTTCCAGGAGTTCTTCGGGTCA 720  
 DB 661 ACATATGTCAGCTGCGGAGCATTAATCACTTGTGCTTCCAGGAGTTCTTCGGGTCA 720  
 QY 721 GATGAAACGACGTATGCTCTTCG 744  
 DB 721 GATGAAACGACGTATGCTCTTCG 744

DB 721 GATGAAACGACGTATGCTCTTCG 744  
 RESULT 2  
 AEB48730  
 ID AEB48730 standard; DNA; 744 BP.  
 XX  
 AC AEB48730;  
 XX  
 DT 22-SEP-2005 (first entry)  
 XX  
 DE *Allicyclobacillus* sp. DSM 15716 functional polypeptide coding sequence.  
 KW Protein production; protein secretion; surfactant; feedstuff; food; gene;  
 ds.  
 XX  
 OS *Allicyclobacillus* sp.; 'DSM 15716'.  
 XX  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..744  
 FT /+tag= b  
 FT /transl\_except= (pos:1..3,aa:Met)  
 FT /partial  
 FT /note= "No stop codon"  
 FT sig\_peptide 1..123  
 FT /+tag= a  
 FT 124..744  
 FT mat\_peptide /+tag= c  
 FT /product= "Mature polypeptide"  
 PN WO200506339-A2.  
 XX  
 PD 21-JUL-2005.  
 XX  
 PF 06-JAN-2005; 2005WO-DK000004.  
 XX  
 PR 06-JAN-2004; 2004DK-00000010.  
 PR 04-FEB-2004; 2004DK-00000165.  
 PR 23-FEB-2004; 2004US-00784592.  
 PR 25-FEB-2004; 2004DK-00000298.  
 XX  
 PA (NOVO ) NOVOZYMES AS.  
 XX  
 PI Wiltling R, Lassen SF, Ostergaard PR;  
 XX  
 DR WP1; 2005-506869/51.  
 DR P-PSDB; AEB48755.  
 XX  
 PT New mature functional polypeptide of *Allicyclobacillus* sp., used for  
 PT commercial research purposes, specifically for cleaning a cellulosic  
 PT fabric, preparing food or feed additive, and for treating lignocellulosic  
 PT materials and pulp.  
 PS Claim 3; SEQ ID NO 18; 151bp; English.  
 XX  
 CC The present invention is based on the finding of a strain of  
 CC *Allicyclobacillus*, namely *Allicyclobacillus* sp. DSM 15716, which grows at  
 CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It  
 CC is an object of the invention to identify and provide polypeptides  
 CC secreted from *Allicyclobacillus* sp. DSM 15716 because such polypeptides  
 CC may be used for industrial purposes and may also be produced in  
 CC industrially relevant processes and amounts. Thus, the invention provides  
 CC isolated mature functional polypeptides which are at least 90% identical  
 CC to, and exhibit and same function as a corresponding secreted polypeptide  
 CC AEB48730-AEB48762 from *Allicyclobacillus* sp. DSM 15716, and also provides  
 CC polynucleotides AEB48733-AEB48737 encoding such polypeptides. The  
 CC polypeptides are preferably enzymes having acid endoglucanase, acid  
 CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl  
 CC protease, serine protease, HtrA-like serine protease, disulfide  
 CC isomerase, gamma-D-glutamyl-L-amino acid endopeptidase, endo-beta-N-  
 CC acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase,  
 CC phytase, phospholipase C, polysaccharide deacetylase, xylanase  
 CC deacetylase or sulfite oxidase activity. A composition comprising such an

CC enzyme can be used in a detergent composition, or a food or feed  
CC composition. Also claimed are: nucleic acid constructs, recombinant  
CC expression vectors and host cells, methods of producing the polypeptides;  
CC a storage medium suitable for use in an electronic device comprising  
CC information of the amino acid sequence of a polypeptide of the invention  
CC or the nucleotide sequence encoding the polypeptide, and a process  
CC comprising employing the polypeptide, or polynucleotide encoding it, in  
CC an industrial or household technical process. The present sequence is the  
CC coding sequence of a functional polypeptide of the invention.

XX Sequence 744 BP; 202 A; 159 C; 179 G; 204 T; 0 U; 0 Other;

Query Match 100.0%; Score 744; DB 14; Length 744;

Best Local Similarity 100.0%; Pred. No. 5.7e-237;

Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGCGAATTATGAAAGTTTGGAGTGGATTTGGTACCGTATATCAATGCTGTTTATTCAG 60
DB 1 GTGCGAATTATGAAAGTTTGGAGTGGATTTGGTACCGTATATCAATGCTGTTTATTCAG 60
QY 61 TGGGGGCGGAATGACGAAATTCGCTTTTCCGGTTCAATTGGGGCAATTAATGTCTTC 120
DB 61 TGGGGGCGGAATGACGAAATTCGCTTTTCCGGTTCAATTGGGGCAATTAATGTCTTC 120
QY 121 GCGAACACGGTGTATGATTCGAGGAAACACCGCGGAGCATCAACGGTAACCGCT 180
DB 121 GCGAACACGGTGTATGATTCGAGGAAACACCGCGGAGCATCAACGGTAACCGCT 180
QY 181 ACAACTCTTTTGGTAAATTCAGAGATAGTTCAACAGTAGCAAGCAAGCAAAATCTG 240
DB 181 ACAACTCTTTTGGTAAATTCAGAGATAGTTCAACAGTAGCAAGCAAGCAAAATCTG 240
QY 241 AGTAGCTCTCCCGCTATAGTCTACGAACTCATTCGCAATGCGCAATCAAGTCTCT 300
DB 241 AGTAGCTCTCCCGCTATAGTCTACGAACTCATTCGCAATGCGCAATCAAGTCTCT 300
QY 301 ACGACTTCATCTTCAGTCGAGGATTAAGATATATCCGTTTCAACATACCGGAAGGTA 360
DB 301 ACGACTTCATCTTCAGTCGAGGATTAAGATATATCCGTTTCAACATACCGGAAGGTA 360
QY 361 GGAGACTTGGAAATTAAGATTAATCCCTGACGACGAAGTTAAGTGTGGGTACGACGG 420
DB 361 GGAGACTTGGAAATTAAGATTAATCCCTGACGACGAAGTTAAGTGTGGGTACGACGG 420
QY 421 ATAGGTGAACCGCAAAATGTCGCTTTGGGTTATCAACATCAACATTAAGAAATGACG 480
DB 421 ATAGGTGAACCGCAAAATGTCGCTTTGGGTTATCAACATCAACATTAAGAAATGACG 480
QY 481 TCCACTCCATGAGAGTGTGATGAGCAATTCATTGTCAGAACTTAAACGGGAACGTT 540
DB 481 TCCACTCCATGAGAGTGTGATGAGCAATTCATTGTCAGAACTTAAACGGGAACGTT 540
QY 541 TATCAGCCGGATTTACTGCTGAGATATATGCAAAATTAATTCAGGGACTATTCGACCC 600
DB 541 TATCAGCCGGATTTACTGCTGAGATATATGCAAAATTAATTCAGGGACTATTCGACCC 600
QY 601 GACCTCAACCCCTGTGTGTCTCATGACGAAATTCGATTTGATATGCGGATTTTATG 660
DB 601 GACCTCAACCCCTGTGTGTCTCATGACGAAATTCGATTTGATATGCGGATTTTATG 660
QY 661 ACATATGTCAGTGTGGGAGATTAATCACTTGTGCTTCATAGGTTTCTTCCGGTCA 720
DB 661 ACATATGTCAGTGTGGGAGATTAATCACTTGTGCTTCATAGGTTTCTTCCGGTCA 720
QY 721 GATGAAGACGATGATGCTCTTCG 744
DB 721 GATGAAGACGATGATGCTCTTCG 744
```

RESULT 3  
AEB45582  
ID AEB45582 standard; DNA; 621 BP.  
XX

```
AC AEB45582;
XX
XX 22-SEP-2005 (first entry)
XX
XX Alicyclobacillus sp. mature functional polypeptide DNA (bases 124-744).
DE
XX Feedstuff; food; detergent; surfactant; pulp; functional polypeptide;
KW gene; ds.
XX
XX Alicyclobacillus sp.; DSM 15716.
OS
XX
XX Key Location/Qualifiers
FH 1.621
FT CDS /tag= a
FT /product= "Alicyclobacillus sp. mature functional
FT polypeptide"
FT /partial
FT /note= "No start and stop codons"
XX
XX US2005147983-A1.
XX
XX 07-JUL-2005.
XX
XX 23-FEB-2004; 2004US-00784592.
XX
XX 06-JAN-2004; 2004DK-00000010.
XX 04-FEB-2004; 2004DK-00000165.
XX
XX (NOVO ) NOVOZYMES AS.
XX
XX Wiltling R, Lassen SF, Ostergaard PR.
XX
XX WPI; 2005-511773/52.
XX
XX P-PSDB; AEB45583.
XX
XX New functional polypeptides having function and amino acid sequence
PT similar to known specific bacterial enzymes useful in industrial,
PT research and household applications e.g. detergents and food.
XX
XX
XX Claim 36; Page; 83pp; English.
XX
XX The present invention relates to functional polypeptides encoded by
CC polynucleotides comprised in the genome of Alicyclobacillus sp. DSM
CC 15716. The invention is useful in industrial, research and household
CC processes such as in composition to prepare feed and food, in detergent
CC formulations and for treating lignocellulosic fabric and pulp. The
CC present sequence is Alicyclobacillus sp. DSM 15716 mature functional
CC polypeptide DNA. Note: This sequence is not shown in the specification
CC but is constructed based on the nucleotide positions provided in claim 36
CC of the specification.
XX
XX
SQ Sequence 621 BP; 177 A; 142 C; 143 G; 159 T; 0 U; 0 Other;
```

Query Match 83.5%; Score 621; DB 14; Length 621;  
Best Local Similarity 100.0%; Pred. No. 5.4e-136;

Matches 621; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 124 AACACGGTATATGATTCGAGAAACACACCGCGAAGCATCAACGGTAAGCGCTACA 183
DB 1 AACACGGTATATGATTCGAGAAACACACCGCGAAGCATCAACGGTAAGCGCTACA 60
QY 184 ACTCTTTGGTAAATTCGAGAAATGTTTCACAGTAGCAAGCAAGCAAAATCTCGAGT 243
DB 184 ACTCTTTGGTAAATTCGAGAAATGTTTCACAGTAGCAAGCAAGCAAAATCTCGAGT 120
QY 244 ACGTCTCCGGCTCAATTAAGTTCACGAACTATTGCAACATGCGCAACATCAAGCTCTACG 303
DB 244 ACGTCTCCGGCTCAATTAAGTTCACGAACTATTGCAACATGCGCAACATCAAGCTCTACG 180
QY 304 ACTTCATCTTCTCAGTCGAAAGTACGATATATCCGTTTCAACATACGGAAGGTAGGA 363
DB 304 ACTTCATCTTCTCAGTCGAAAGTACGATATATCCGTTTCAACATACGGAAGGTAGGA 240
```

QY 364 GACTTGAATTAAGATTAACTCCTGAGCAAGTTAAGTGGGGTACGACCGGATTA 423  
DB 241 GACTTGAATTAAGATTAACTCCTGAGCAAGTTAAGTGGGGTACGACCGGATTA 300  
QY 424 GGTGAACCGCAATATGTCGTTTGGTTATCAACATCAATCAATGAAGATGATCC 483  
DB 301 GGTGAACCGCAATATGTCGTTTGGTTATCAACATCAATGAAGATGATCC 360  
QY 484 ACTCCATGAGAGTGGTATGATGAGCATATTCATTTGAGCACTTAAGCGGAAGTTAT 543  
DB 361 ACTCCATGAGAGTGGTATGATGAGCATATTCATTTGAGCACTTAAGCGGAAGTTAT 420  
QY 544 CAGCCGAGATTTACTGCTGAGATATATGCAAAATTAACAAGGATTAATCCGAC 603  
DB 421 CAGCCGAGATTTACTGCTGAGATATATGCAAAATTAACAAGGATTAATCCGAC 480  
QY 604 CTCACACCTGCTGTGTCCATGACGACAAATCTGCTATTTGATATGCGGATTTATGACA 663  
DB 481 CTCACACCTGCTGTGTCCATGACGACAAATCTGCTATTTGATATGCGGATTTATGACA 540  
QY 664 TATGCTACGTCGGGAGCATTAATCACTGTGCGCTTCATGAGGTTTCTTGGGGTCAAT 723  
DB 541 TATGCTACGTCGGGAGCATTAATCACTGTGCGCTTCATGAGGTTTCTTGGGGTCAAT 600  
QY 724 GAAACGACGTATGCTCTTCCG 744  
DB 601 GAAACGACGTATGCTCTTCCG 621

RESULT 4  
ADA71938/c  
ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; da.

OS Oryza sativa.

XX MO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001MO-1B001105.

XX 22-JUN-2001; 2001MO-1B001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,  
XX Katagiri F, Quan Y, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;  
XX MPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
XX PT pathogenic infection for conferring resistance or tolerance to a plant to  
XX PT bacterial, fungal or viral infection by determining or detecting plant  
XX PT gene expression.

XX Claim 27; SEQ ID NO 5263; 899bp; English.

XX The present invention relates to a method (MI) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. MI  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX CC expression of the gene in an uninfected plant, in a mutant plant that  
XX CC does not express a gene associated with response to pathogenic infection,  
XX CC or in a corresponding incompatible or compatible interaction. (MI) is

CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 5.3%; Score 39.4; DB 8; Length 2000;  
Best Local Similarity 9.1%; Pred. No. 0.086;  
Matches 61; Conservative 304; Mismatches 300; Indels 4; Gaps 2;

QY 24 ATGATTTTGGTACCGTATATCATGCTTTTATTCAGTGGGGCGAATGAACAGATTTCT 83  
DB 1002 MTKYTKYMTATYSSFWKMYAYKWRVAMSSRKTYWCITGGKRYATYCGTKMAAGRW 943  
QY 84 GCGTTTCCGGTTCATTTGTTGGGCAATTAATGTTCTTCCGCAACACGGTATATGATTCG 143  
DB 942 RNAMCWYCCMMKMMWMTSCMMWKRTYSCTYTMGMARVAYAMRRRRYTKMSWRM 883  
QY 144 AGAACAACACCGCGGAACGATCAACGCTAAGCGCTAACCTTTGTTAATTCGAC 203  
DB 882 YMTMTKMAWMTTMCMAKMYATATGMATMMWRRYMYCYAMTCAKCKYMAWMTKMTT 823  
QY 204 GAATATGTTCAAGTAGCAAGCAAGCAAAATTCGAGTACGTCTCCGCTCATATGC 263  
DB 822 WACAWRATSWBRBAAGRWKRYKRYKRYWWRWWRWVAGARWVKRYR--WTKRYA 766  
QY 264 TACGAACATTTGCAACATGCGCAACATCAAGCTGCTACGATCTTCTCACTGCA 323  
DB 765 TRYYKMAWMTWMSWRKWSRMSGMRMSAMRYCSRKCAKTYASARTKAK 706  
QY 324 GTTACATATATCCCGTTTCAACATACGGAAGTAGAGACTTGAATTAAGTTA 383  
DB 705 RSYRFRRRMYWKMTYRMYRMSCRMTBAMSRBRMAGASMSCMWYRGAWSMYS 646  
QY 384 CTCCTGCAAGCAAGTTAAGAGTGGGTAGACGCGATAGGTAAACCGCAATGTCG 443  
DB 645 KYSCAKCKKRYWTSYSMTGMYSYKSMWTSKMSYMGMTCTMTYSMSKSTR 586  
QY 444 GTTTGGTTATCAACATCAACATTAAGAAAGAGGATCCACTCTATGAGAGGTGTA 503  
DB 585 SMGRMSGMSRMYWMMKMRKRYMYKMYCTWRRCCTRMGTMTTTSBRMMYGR 526  
QY 504 TGGCATAT-TCCATTTGAGAACTTAACGGAACGTTATTCAGCCGATTTACTGCTG 562  
DB 525 YKARTYSRRYMYKYRYCYMYTYGMYMKOSYMMRYCYCAKCKKCYAMCKAAYSGMM 466  
QY 563 AGATATATGCAAAATTAATTCAGGACTATTTCCAGCAGACTCAACCTGTGTGCA 622  
DB 465 MYWRYKYSKWRMSYKMYMSWYKCRSMKYGAKGCGCKMYTCSYGMKYTYTGSYK 406  
QY 623 TGACGCAAAATCTGTAATTTGATATGCGGATTTTATGACATATGTCACGTCGGGACG 682  
DB 405 YSRCKYKMYMYKMMYMMYTSAYSMWYTYAKTKYKRGYKRGYKRGYKRGYKRGY 346  
QY 683 ATACTCAC 691  
DB 345 CTWMCYMKC 337

RESULT 5  
ADA71938  
ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX KW gene; da.

OS Oryza sativa.

```
XX XX MO2003000898-A1.
XX PN
XX PD 03-JAN-2003.
XX PF 22-JUN-2001; 2001WO-IB001105.
XX PR 22-JUN-2001; 2001WO-IB001105.
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX DR MPI; 2003-175290/17.
XX PT Identifying at least one gene involved in plant resistance or response to
XX PT pathogenic infection for conferring resistance or tolerance to a plant to
XX PT bacterial, fungal or viral infection by determining or detecting plant
XX PT gene expression.
XX PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX CC The present invention relates to a method (M1) for identifying genes
XX CC involved in plant resistance or response to pathogenic infection. M1
XX CC comprises identifying a gene whose expression is significantly altered in
XX CC the incompatible interaction of plant gene expression relative to
XX CC expression of the gene in an uninfected plant, in a mutant plant that
XX CC does not express a gene associated with response to pathogenic infection,
XX CC or in a corresponding incompatible or compatible interaction. (M1) is
XX CC useful for conferring resistance to resistance or tolerance to a plant to
XX CC bacterial, fungal or viral infection. The present sequence was used to
XX CC illustrate the invention.
XX SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match      4.9%; Score 36.6; DB 8; Length 2000;
Best Local Similarity 8.2%; Pred. NO. 0.73; Mismatches 249; Indels 7; Gaps 3;
Matches 49; Conservative 290;

QY 82 CTGCGTTTGGCCGTTTCATTTGGGCAATTAATTTCTTCCGCAACGCGTATATGATT 141
DB 256 SRRBSMMWMMKMRKRSYGMYSWYKMMCTAYKKSYSRKYRGGCGRATRYVG 315
QY 142 CGAGAAACACACCGCGAAGCATCAACGCTAACGCTTCTT-----TGGTTA 196
DB 316 RGYMSRMMWMMYKMYRKYKMGKMGMMRMMRSMCRMSKACCYMRMRMRMTERRRM 375
QY 197 ATTGACGCAATAGTTCCACAGTAGAAGCAAGCAAGCAAACTCGATAGTCTCCCGCTC 256
DB 376 AKKSSPTSRKRRKRWCMRRKRYKRMGYSRMSCKRARMKCRSGRARMKCGRGMTGR 435
QY 257 ATTAGCTCTACGA-ACCTGATGCAATGCGCAACATCACTGCTGCTCACTTCTTCT 315
DB 436 MTSYGMRRKMSKMRKASKTKMMSRMRYMRKCKSRITTMGKTRGMMGMRKRYKRG 495
QY 316 CAGTCGAAGTAC-GATATATCCCGTTTCAACATACGGAAGAGTAGAGACTTGGAAAT 374
DB 496 MKRKCRRRRWGMWRMRMYRSAYTYRCAKRYKYSARAKRKACWYRKGXYWAGMW 555
QY 375 TAGAGTTAATCTCCCTGACAGCAAGTTAAGAGTGTGGGTGCGACGCGGATAGTGAACCGC 434
DB 556 MRRYKRMYYKMMWYKRYKSKSWYCKMSYASCKMSKRGAKRKCKRMSKMSKMSR 615
QY 435 AAATGTCGTTTGGGTATATCAACATCAACATTAAGAATGACGATCCACTATGGA 494
DB 616 SRRKCKKSKSKSSARBYAMMGMSGMSRMSKSTYCTWRKMSKSKSTCTMYMYMSKT 675
QY 495 GGTCTGTTAGCATATTCATTTGCAAGACTTAAACGGGAAGCTTTATCACCGGATTC 554
DB 676 YAKYGSYMYRYRACWMMYMRWYRYRYSYTYMAATYSTTMAATGMKYSRYWTSYVK 735
QY 555 TACTGTGAGATATATGCAAAATCAAAATTCAGGAGACTATTCGACCGAAGCTCAACCTCG 614
```

```
DB 736 YCKCSWKRSMMYMSWMAKTKWKKRRYATMMWMMYYSKMYTWCTMMGWTMYWK 795
QY 615 TGTGTCATGACGACAAATCTGTATTGATATGCCGATTTTATGACATATGCT 669
DB 796 YMKRMYMYKCTKYTWYMGATYTWTGWAAMWMAKTKMRMGMTGAKTRGBARKRY 850

RESULT 6
ACCA47595
ID ACC47595 standard; DNA; 2306 BP.
AC ACC47595;
XX 23-JUN-2003 (first entry)
XX DE Gastrodia elata antifungal protein GAFP-encoding gene.
XX KW Gastrodia antifungal protein; GAFP; fungus-inducible promoter; plant;
XX KW transgenic; gene; ds.
XX OS Gastrodia elata.
XX FH Key location/Qualifiers
XX FT misc_feature 1..1822
XX FT /tag= a
XX FT /note= "This region is specifically claimed in claim 1"
XX FT promoter 1..1305
XX FT /tag= b
XX FT /note= "Specifically claimed in claim 2"
XX FT CDS 1307..1822
XX FT /tag= c
XX FT /product= "GAFP"
XX FT /note= "Gastrodia antifungal protein"
XX PN CN136045-A.
XX PD 28-AUG-2002.
XX PF 16-JAN-2001; 2001CN-00100478.
XX PR 16-JAN-2001; 2001CN-00100478.
XX PA (GENE-) INST GENETICS CHINESE ACAD SCI.
XX PI Sun Y, Sha Q, Li W;
XX PI MPI; 2003-240396/24.
XX DR Inducible promoter of fungus.
XX PT Inducible promoter of fungus.
XX PS Claim 1; Page 7-8 (Disclosure); 11pp; Chinese.
XX CC The invention relates to a fungus-inducible promoter from the plant
XX CC Gastrodia elata. The promoter is that of the Gastrodia antifungal protein
XX CC (GAFP) gene which is expressed in the tubers of the plant. The fungus-
XX CC inducible promoter may be used in transgenic plants to mediate expression
XX CC of heterologous genes in the presence of fungus. The present sequence
XX CC represents the GAFP gene sequence which includes its promoter
XX SQ Sequence 2306 BP; 735 A; 416 C; 441 G; 714 T; 0 U; 0 Other;

Query Match      4.9%; Score 36.2; DB 8; Length 2306;
Best Local Similarity 57.5%; Pred. NO. 1.1; Mismatches 48; Indels 0; Gaps 0;
Matches 65; Conservative 0;

QY 379 GTTAATCTCCCTGACAGCAAGTTAAGGTGTGGGCTACGACGCGATGAGTGAACCGCAAT 438
DB 1631 GGTAACTACTATCTATCTATCTTCAAGAGATGTAACGTCCTATATGATATTTCTAT 1690
QY 439 GGTGCGTTTGGGTATATCAACATCAACATTAAGAATGACGAGTCCACTCTCTAT 491
DB 1691 AATGCAATTTGGGCAACCAACCAACGATGGAATCTGAAATATCATCTCTAT 1743
```

```
RESULT 7
ADA13411.1
Continuation (2 of 4) of ADA13411 from base 100001 (Human transporter protein genomic DN
WP Sequence split into 4 fragments LOCUS ADA13411 Accession Ada13411
WP Fragment Name Begin End
WP ADA13411_0 1 110000
WP ADA13411_1 100001 210000
WP ADA13411_2 200001 310000
WP ADA13411_3 300001 389627

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 9; Length 110000;
Matches 62; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 102 GGGGCAATTAATGTTCTTCGCAACGCGTATATGATTGAGGAAACACCGCGGAA 161
Db 56226 GAGGGACAGATTCCTTCTACAAAGCCATGTTGTGTTCTTGAAAGACACGCAAAA 56285

Qy 162 CGCATCAACGCTAAGCGCTACAACCTTCTTGTTAATTCGACGAA 206
Db 56286 CTTTCAAGGCAAACTGTTTCATTCATTCGGATTAACCAAGAA 56330

RESULT 8
ADO59446.1
Continuation (2 of 4) of ADO59446 from base 100001 (Human cancer-associated (CA) gene se
WP Sequence split into 4 fragments LOCUS ADO59446 Accession Adq59446
WP Fragment Name Begin End
WP ADO59446_0 1 110000
WP ADO59446_1 100001 210000
WP ADO59446_2 200001 310000
WP ADO59446_3 300001 405660

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 12; Length 110000;
Matches 62; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 102 GGGGCAATTAATGTTCTTCGCAACGCGTATATGATTGAGGAAACACCGCGGAA 161
Db 64226 GAGGGACAGATTCCTTCTACAAAGCCATGTTGTGTTCTTGAAAGACACGCAAAA 64285

Qy 162 CGCATCAACGCTAAGCGCTACAACCTTCTTGTTAATTCGACGAA 206
Db 64286 CTTTCAAGGCAAACTGTTTCATTCATTCGGATTAACCAAGAA 64330

RESULT 9
ADZ13757.1
Continuation (2 of 4) of ADZ13757 from base 100001 (Human cancer-associated genomic DNA
WP Sequence split into 4 fragments LOCUS ADZ13757 Accession Adz13757
WP Fragment Name Begin End
WP ADZ13757_0 1 110000
WP ADZ13757_1 100001 210000
WP ADZ13757_2 200001 310000
WP ADZ13757_3 300001 408438

Query Match
Best Local Similarity 4.9%; Score 36.2; DB 14; Length 110000;
Matches 62; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 102 GGGGCAATTAATGTTCTTCGCAACGCGTATATGATTGAGGAAACACCGCGGAA 161
Db 64474 GAGGGACAGATTCCTTCTACAAAGCCATGTTGTGTTCTTGAAAGACACGCAAAA 64533

Qy 162 CGCATCAACGCTAAGCGCTACAACCTTCTTGTTAATTCGACGAA 206
Db 64534 CTTTCAAGGCAAACTGTTTCATTCATTCGGATTAACCAAGAA 64578

RESULT 10
ABAI7217
#
```

```
ID ABAI7217 standard; DNA; 7636 BP.
XX
AC ABAI7217;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 9548.
XX
KW Human; nocitropic; neuroprotective; cytosatic; dermatological; virocidic;
KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnecacy;
KW antiparkinsonian; antisticking; antianaemic; antirheumatic; cancer;
KW antirheumatic; hepatocytropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0188874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227183P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
```







OY		475	GACGATCCACTCCTTAATGAGTGTGAATGCATATTTCATTTCAGAACTTAACGGG	534
Dd		1390	AAGGATCTAATCGTTAATGAGATGAGTAGTAATGACGAATGCTCAAACTAAACCCA	1449
OY		535	AAC 537	
Db		1450	AAC 1452	
RESULT 13				
ID	AAK6772			
XX	AAK6772 standard; DNA; 42488 BP.			
XX				
AC	AAK6772;	.		
XX				
DT	06-NOV-2001	(first entry)		
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21584.			
XX				
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;			
KM	cytotoxic; gene therapy; vaccine; metastasis; ds.			
XX				
OS	Homo sapiens.			
PN	WO200157182-A2.			
XX				
PD	09-AUG-2001.			
XX				
Pf	17-JAN-2001;	2001WO-US001354.		
XX				
PR	31-JAN-2000;	2000US-0179065P.		
PR	04-FEB-2000;	2000US-0180628P.		
PR	24-FEB-2000;	2000US-0184664P.		
PR	02-MAR-2000;	2000US-0186350P.		
PR	16-MAR-2000;	2000US-0189874P.		
PR	17-MAR-2000;	2000US-0190076P.		
PR	18-APR-2000;	2000US-0198123P.		
PR	19-MAY-2000;	2000US-0205151P.		
PR	07-JUN-2000;	2000US-0209467P.		
PR	28-JUN-2000;	2000US-0214886P.		
PR	30-JUN-2000;	2000US-0215135P.		
PR	07-JUL-2000;	2000US-0216647P.		
PR	07-JUL-2000;	2000US-0216880P.		
PR	11-JUL-2000;	2000US-0217487P.		
PR	11-JUL-2000;	2000US-0217496P.		
PR	14-JUL-2000;	2000US-0218290P.		
PR	26-JUL-2000;	2000US-0220963P.		
PR	26-JUL-2000;	2000US-0220964P.		
PR	14-AUG-2000;	2000US-0224518P.		
PR	14-AUG-2000;	2000US-0224519P.		
PR	14-AUG-2000;	2000US-0225213P.		
PR	14-AUG-2000;	2000US-0225214P.		
PR	14-AUG-2000;	2000US-0225266P.		
PR	14-AUG-2000;	2000US-0225267P.		
PR	14-AUG-2000;	2000US-0225267P.		
PR	14-AUG-2000;	2000US-0225270P.		
PR	14-AUG-2000;	2000US-0225270P.		
PR	14-AUG-2000;	2000US-0225447P.		
PR	14-AUG-2000;	2000US-0225575P.		
PR	14-AUG-2000;	2000US-0225758P.		
PR	14-AUG-2000;	2000US-0225759P.		
PR	18-AUG-2000;	2000US-0226279P.		
PR	22-AUG-2000;	2000US-0226681P.		
PR	22-AUG-2000;	2000US-0226686P.		
PR	22-AUG-2000;	2000US-0227182P.		
PR	23-AUG-2000;	2000US-0227009P.		
PR	30-AUG-2000;	2000US-0228924P.		
PR	01-SEP-2000;	2000US-0229287P.		
PR	01-SEP-2000;	2000US-0229343P.		
PR	01-SEP-2000;	2000US-0229344P.		
PR	01-SEP-2000;	2000US-0229345P.		
PR	05-SEP-2000;	2000US-0229509P.		
PR	05-SEP-2000;	2000US-0229513P.		
PR	06-SEP-2000;	2000US-0230437P.		

PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231142P
PR	08-SEP-2000	2000US-0231143P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0232081P
PR	08-SEP-2000	2000US-0232080P
PR	12-SEP-2000	2000US-0231368P
PR	14-SEP-2000	2000US-0232397P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232399P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0232403P
PR	14-SEP-2000	2000US-0233063P
PR	14-SEP-2000	2000US-0233064P
PR	14-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234423P
PR	21-SEP-2000	2000US-0234424P
PR	21-SEP-2000	2000US-0234597P
PR	25-SEP-2000	2000US-0234978P
PR	25-SEP-2000	2000US-0234988P
PR	25-SEP-2000	2000US-0235484P
PR	27-SEP-2000	2000US-0235853P
PR	27-SEP-2000	2000US-0235854P
PR	27-SEP-2000	2000US-0235857P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236367P
PR	29-SEP-2000	2000US-0236368P
PR	29-SEP-2000	2000US-0236369P
PR	29-SEP-2000	2000US-0236370P
PR	02-OCT-2000	2000US-0236802P
PR	02-OCT-2000	2000US-0237037P
PR	02-OCT-2000	2000US-0237038P
PR	02-OCT-2000	2000US-0237040P
PR	02-OCT-2000	2000US-0237043P
PR	13-OCT-2000	2000US-0239935P
PR	13-OCT-2000	2000US-0240963P
PR	20-OCT-2000	2000US-0241808P
PR	20-OCT-2000	2000US-0241809P
PR	20-OCT-2000	2000US-0241821P
PR	20-OCT-2000	2000US-0241822P
PR	20-OCT-2000	2000US-0241786P
PR	20-OCT-2000	2000US-0241878P
PR	20-OCT-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246476P
PR	08-NOV-2000	2000US-0246477P
PR	08-NOV-2000	2000US-0246478P
PR	08-NOV-2000	2000US-0246523P
PR	08-NOV-2000	2000US-0246524P
PR	08-NOV-2000	2000US-0246525P
PR	08-NOV-2000	2000US-0246526P
PR	08-NOV-2000	2000US-0246527P
PR	08-NOV-2000	2000US-0246528P
PR	08-NOV-2000	2000US-0246532P
PR	08-NOV-2000	2000US-0246609P
PR	08-NOV-2000	2000US-0246610P
PR	08-NOV-2000	2000US-0246611P
PR	17-NOV-2000	2000US-0246613P
PR	17-NOV-2000	2000US-0249077P
PR	17-NOV-2000	2000US-0249208P
PR	17-NOV-2000	2000US-0249209P
PR	17-NOV-2000	2000US-0249210P
PR	17-NOV-2000	2000US-0249211P
PR	17-NOV-2000	2000US-0249212P
PR	17-NOV-2000	2000US-0249213P
PR	17-NOV-2000	2000US-0249214P
PR	17-NOV-2000	2000US-0249215P
PR	17-NOV-2000	2000US-0249216P
PR	17-NOV-2000	2000US-0249217P
PR	17-NOV-2000	2000US-0249218P



XX	RESULT 15
ADXL12170	ID
ADXL12170 standard; cDNA; 729 BP.	
XX	AC
ADXL12170;	
XX	DT
21-APR-2005 (first entry)	
XX	DE
Plant full length insert polynucleotide seqid 6745.	
KW	plant protectant; plant growth regulant; gene therapy; plant;
KW	recombinant DNA construct; physical array; plant breeding marker;
KW	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW	extreme osmotic condition; pathogens tolerance; pest tolerance;
KW	growth rate; cell cycle pathway; disease resistance;
KW	galactomannan production; lignin production; plant growth regulator;
KW	yield; plant growth; plant development; seed oil; protein yield;
XX	protein content; gene; ss.
OS	Unidentified.
XX	
PN	US2004034888-A1.
PD	19-FEB-2004.
XX	
PF	28-APR-2003; 2003US-00425114.
XX	
PR	06-MAY-1999; 99US-00304517.
PR	05-NOV-2001; 2001US-00985678.
XX	
PA	(LTUJ/) LTU J.
PA	(ZHOU/) ZHOU Y.
PA	(KOVA/) KOVALIC D K.
PA	(SCRE/) SCREEN S E.
PA	(TABAA/) TABAASKA J B.
PA	(CAOY/) CAO Y.
XX	
PI	Liu J, Zhou Y, Kovacic DK, Screen SE, Tabaska JB, Cao Y;
DR	WPI; 2004-180133/17.
XX	
PT	New recombinant DNA construct, useful for improving plant tolerance to
PT	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT	pests, for conferring increased resistance to plant disease, or for
PT	improving yield.
XX	
PS	Claim 1; SEQ ID NO 6745; 15bp; English.
CC	
CC	The invention describes a recombinant DNA construct comprising a
CC	polynucleotide consisting of a sequence encoding an amino acid sequence
CC	available in electronic form from the US patent office at
CC	ftp.segdata.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
CC	of the invention are also useful in physical arrays of molecules and as
CC	plant breeding markers. The recombinant DNA construct is useful for
CC	improving plant tolerance to cold, heat, drought, herbicides, extreme
CC	osmotic conditions, pathogens or pests, for manipulating growth rate in
CC	plant cells by modification of the cell cycle pathway, for conferring
CC	increased resistance to plant disease, for producing galactomannan,
CC	lignin or plant growth regulators, for increasing the rate of homologous
CC	recombination in plants, for improving yield by modification of
CC	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC	or by providing improved plant growth and development under at least one
CC	stress condition or for modifying seed oil or protein yield and/or
CC	content. This sequence represents a plant full length insert
CC	polynucleotide that can be used in the recombinant DNA construct of the
CC	invention.
XX	
SO	Sequence 729 BP; 208 A; 152 C; 171 G; 198 T; 0 U; 0 Other;
Query Match	4.5%; Score 33.8; DB 13; Length 729;
Best Local Similarity	52.5%; Pred. No. 3.9;
Matches 74; Conservative	0; Mismatches 67; Indels 0; Gaps 0

Oy	584	CAGGGACATATTTCCACCGACCTCAACCTCGTGTGTCCATGACGACAAATCTGTATTG	643
Db	464	CTGGGCATACATGCACCGAGTACAAAGCCACCGAGTGCACAAAGATTTAGGAAATGGG	523
Oy	644	ATATGCCCGATTTTATGACATATATGTCACGTGGGACGATTACTCATTTGTGCTTCCA	703
Db	524	TTTCAACGGGATCTCTCTGATATATTAACGTGCTGAGAGATTAACGTACGTGTCTCTTA	583
Oy	704	TGGGTTTCTTGGGTGAGATG	724
Db	584	TGAGTGATGGCCGACGTACG	604

Search completed: April 7, 2006, 21:33:54  
Job time : 597 secs

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 21:26:27 ; Search time 3831 Seconds

(without alignments)  
9086.293 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744  
Sequence: 1 gtcgcaattatgaagttt.....aacgacgtatgcttccg 744

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 2339354128 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lifting first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_hic:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_est7:\*

9: gb\_gsa1:\*

10: gb\_gsa2:\*

11: gb\_gsa3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.8	5.3	444	1	AU269824 AU269824
2	39.6	5.3	714	6	CA996800 r903d03.Y
3	39.2	5.3	475	3	BJ445212 BJ445212
4	39.2	5.3	729	5	BU364247 BU364247
5	39.2	5.3	782	3	BU436673 BU436673
6	39	5.2	393	3	BU439058 BU439058
7	39	5.2	742	8	DR070622 RTDK1.14
8	39	5.2	782	7	CO167205 FLD1.67.C
9	38.6	5.2	374	3	BJ440426 BJ440426
10	38.6	5.2	374	3	BJ440571 BJ440571
11	38.6	5.2	391	3	BU440677 BU440677
12	38.6	5.2	393	3	BU440820 BU440820
13	38.6	5.2	394	3	BU440720 BU440720
14	38.6	5.2	422	3	BU437922 BU437922
15	38.6	5.2	458	3	BU444052 BU444052
16	38.6	5.2	463	3	BU438307 BU438307
17	38.6	5.2	463	3	BU445358 BU445358
18	38.6	5.2	470	3	BU438852 BU438852
19	38.6	5.2	472	3	BU437919 BU437919
20	38.6	5.2	503	3	BU444250 BU444250
21	38.6	5.2	506	3	BU442298 BU442298
22	38.6	5.2	524	3	BU443637 BU443637

C	23	38.6	5.2	526	3	BJ441822	BJ441822
C	24	38.6	5.2	531	1	AU268364	AU268364
C	25	38.6	5.2	532	3	BU439175	BU439175
C	26	38.6	5.2	535	3	BU442262	BU442262
C	27	38.6	5.2	539	3	BU444947	BU444947
C	28	38.6	5.2	541	3	BU444725	BU444725
C	29	38.6	5.2	543	3	BU446010	BU446010
C	30	38.6	5.2	547	3	BU437840	BU437840
C	31	38.6	5.2	555	3	BU444817	BU444817
C	32	38.6	5.2	557	3	BU442366	BU442366
C	33	38.6	5.2	558	3	BU443146	BU443146
C	34	38.6	5.2	558	3	BU446802	BU446802
C	35	38.6	5.2	567	3	BU437291	BU437291
C	36	38.6	5.2	570	3	BU442058	BU442058
C	37	38.6	5.2	571	3	BU437532	BU437532
C	38	38.6	5.2	572	3	BU442947	BU442947
C	39	38.6	5.2	573	3	BU445870	BU445870
C	40	38.6	5.2	575	3	BU446620	BU446620
C	41	38.6	5.2	578	3	BU436557	BU436557
C	42	38.6	5.2	578	3	BU438536	BU438536
C	43	38.6	5.2	578	3	BU444484	BU444484
C	44	38.6	5.2	579	3	BU442001	BU442001
C	45	38.6	5.2	580	3	BU446126	BU446126

#### ALIGNMENTS

RESULT 1  
AU269824 444 bp mRNA linear EST 26-APR-2004

LOCUS AU269824 VS Dictyostelium discoideum cDNA clone VSJ372 3', mRNA

DEFINITION AU269824 sequence.

ACCESSION AU269824

VERSION AU269824.1 GI:20528622

KEYWORDS EST

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Dictyostelium discoideum

AUTHORS 1 (bases 1 to 444)

Umeshihara,H., Morio,T., Saito,T., Kohara,Y., Koriiki,E., Ochiai,H.,

Maeda,M., Williams,J.G., Takeuchi,I., and Tanaka,Y.

Analyses of cDNAs from growth and slug stages of Dictyostelium

discoideum

Nucleic Acids Res. 32 (5), 1647-1653 (2004)

JOURNAL 15010511

PUBMED Contact: Hideko Umeshihara

COMMENT Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp.

location/Qualifiers

1..444

/organism="Dictyostelium discoideum"

/mol\_type="mRNA"

/strain="AX4"

/db\_xref="taxon:44689"

/clone="VSJ372"

/sex="mat A"

/dev\_stage="vegetative"

/clone\_lib="VS"

ORIGIN

Query Match 5.3%; Score 39.8; DB 1; Length 444;

Best Local Similarity 51.4%; Pred. No.0.51;

Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

422 TAGGTGAACCGCAATGTCGTTTGGTTATCAATCAATCAATGAATGACGAT 481  
DB 70 TGGTTGAAGAAAGAAATTAAGTATGATATCAACAAACCGTATTATATGCGAAT 129

QY 482 CCATCTCATGAGGTCGTGATGATTCATTCATTTGACGAAGCTTAAACGGAAAGCTTT 541  
DB 130 CTTTGTGTAATGATTCNCCAGGTCGTGATTCATTCATTAATGGGTAATGTAAGATG 189  
QY 542 ATCAAGCGGATTCATCTGCTGATGATATGCAATATCAATTCAGGACTATT 594  
DB 190 CAATTCGAATGTTACTGTGTGAAGTTATGCTATTCATTCAAATCTGCTAAATAAT 242

RESULT 2  
CA996800/c 714 bp mRNA linear EST 07-JAN-2003  
LOCUS rg03d03.y1 Meloidogyne hapla J2 pAMP1 v1 Meloidogyne hapla cDNA 5'  
DEFINITION similar to TR:001685 001685 SIMILAR TO ALANINE AMINOTRANSFERASE.  
[1] ; mRNA sequence.  
ACCESSION CA996800.1 GI:27541671  
VERSION EST.  
KEYWORDS Meloidogyne hapla  
ORGANISM Meloidogyne hapla  
SOURCE Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;  
Tylenchoidae; Meloidogynidae; Meloidogyninae; Meloidogyne.  
REFERENCE 1 (bases 1 to 714)  
McCartier,J., Clifton,S., Chapell,J.B., Page,D., Martin,J.,  
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theisling,B.,  
Bowers,Y., Gibbons,M., Ritzer,B., Bennett,J., Franklin,C.,  
Teagarden,Shvill,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,  
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,  
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McGinn,R., Waterson,R. and Wilson,R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
CONTACT: McCarter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was constructed by Claire Murphy and Dr. James McCarter  
at Washington University, St. Louis. J2 were provided by Dr.  
Valerie Williamson of the University of California at Davis  
(vwilliamson@ucdavis.edu).  
Seq primer: -40RP from Gibco  
High quality sequence stop: 414.  
FEATURES  
location/Qualifiers  
1..714  
/organism="Meloidogyne hapla"  
/mol\_type="mRNA"  
/db\_xref="taxon:6305"  
/dev\_stage="J2"  
/lab\_host="DH10B"  
/clone\_lib="Meloidogyne hapla J2 pAMP1 v1"  
/note="Vector: pAMP1 (Gibco) ; Site\_1: NotI; Site\_2: SalI;  
The library was constructed by Claire Murphy and Dr. James  
McCarter at Washington University, St. Louis. The cDNA was  
made by using Dynabead oligo-dT priming (Dynal). PCR based  
library using a modified protocol from the SMART PCR cDNA  
Synthesis Kit from Clontech. Directionally cloned into the  
UDG sites of pAMP1. J2 were provided by Dr. Valerie  
Williamson of the University of California at Davis  
(vwilliamson@ucdavis.edu)."

ORIGIN  
Query Match 5.3%; Score 39.6; DB 6; Length 714;  
Best Local Similarity 50.5%; Pred. No. 0.69;  
Matches 96; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 443 CGTTTGGGTATTCACATCACCATAAGAAAGACGACCTCTATGAGAGTCGTTG 502  
DB 268 CATTGGATTAATCGTTCACACGATGACACGCTTCTGCAAAATTTGAACGTGTTG 209  
QY 503 ATGGCATATTCATTCGACGAAGCTTAAACGGAAAGCTTATGACGCGATTCAGCTG 562

DB 208 TTGACATATTTGCTTATATTCACAATTAATGAAATGTAAGTCTTGTCGCTG 149  
QY 563 AGATATATGCAATATCAATTCAGGACTATTTCCAGCAGCTCAACCTGTGTGTC 622  
DB 148 GAATATTAAGTAAACACAGACAGTCGATTAATGCTGCTCAATCTCAAGCTGTTGAATCA 89  
QY 623 TGACGACAAA 632  
DB 88 TTCAAAAAA 79

RESULT 3  
BU445212/c 475 bp mRNA linear EST 13-MAR-2002  
LOCUS BU445212/c  
DEFINITION BU445212 Dictyostelium discoideum cDNA library, VF Dictyostelium  
discoideum cDNA clone dv58e21 3', mRNA sequence.  
ACCESSION BU445212.1 GI:19419933  
VERSION EST.  
KEYWORDS Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
SOURCE Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
REFERENCE 1 (bases 1 to 475)  
Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-I,T.  
Full length cDNA of Dictyostelium discoideum at the vegetative  
stage  
Unpublished (2002)  
CONTACT: Tadasu Shin-I  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshin@genes.nig.ac.jp.  
FEATURES  
location/Qualifiers  
1..475  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="dv58e21"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN  
Query Match 5.3%; Score 39.2; DB 3; Length 475;  
Best Local Similarity 51.4%; Pred. No. 0.81;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCGCAATGTCGCTTTGGGTTATCAATCACCCTAAGAAATGCGAT 481  
DB 269 TGGTGAAGAAAGAAATATTAAGTTATGATATCAACAAACCTATTANATGCGAAT 210  
QY 482 CCATCTCATGAGGTCGTGATGATTCATTCATTTGACGAAGCTTAAACGGAAAGCTTT 541  
DB 209 CTTTGTGTAATGATACCAAGTCGTTTACCTTACATTAATTAATGGGTTATGTAAGATG 150  
QY 542 ATCAACCGGATTCATCTGCTGATATGCAATATCAATTCAGGACTATT 594  
DB 149 CAATTCGAATGTTACTGTGTGAAGTTATGCTATTCATTCAAATCTGCTAAATAAT 97

RESULT 4  
BU364247/c 729 bp mRNA linear EST 28-NOV-2002  
LOCUS BU364247 603785262F1 CSEQHN72 Gallus gallus cDNA clone CHSST740k15 5', mRNA  
DEFINITION sequence.  
ACCESSION BU364247  
VERSION BU364247.1 GI:25872248  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)



ORIGIN /dev stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

Query Match 5.2%; Score 39; DB 3; Length 393;  
Best Local Similarity 50.0%; Pred. No. 0.88;  
Matches 84; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCGCAATGCGTGGTTGGTATGCAACATCCATTAAGAAATGACGGAT 481  
DB 255 TGGTGAAGAAAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 196  
QY 482 CCATCTATAGAGGTCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 541  
DB 195 CTCTTATGAT 136  
QY 542 ATGACCGGATCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 589  
DB 135 CAATTCNATGTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 88

RESULT 7  
LOCUS DR070622 742 bp mRNA linear EST 08-JUN-2005  
DEFINITION RTDX1\_14\_C03.g1\_A029 Roots, dark Pinus taeda cDNA clone  
LOCUS DR070622 RTDX1\_14\_C03\_A029 5', mRNA sequence.  
ACCESSION DR070622  
VERSION DR070622.1 GI:67048356  
KEYWORDS EST.  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Pinus taeda

REFERENCE Buzaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 742)  
AUTHORS Pratt, L., Cordonnier-Pratt, M.M., Lorenz, W.W., Zimmermann, C. and Dean, J.F.D.  
TITLE An EST database from dark-treated loblolly pine (Pinus taeda) roots unpublished (2005)  
JOURNAL Other ESTs: RTDX1\_14\_C03.b1\_A029  
COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

#### FEATURES

source Location/Qualifiers  
1..742  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="3 CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTDX1\_14\_C03\_A029"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Roots, dark"  
/note="Organ: Root; Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 50 days (May 1 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x

nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. Twenty-four hours (24h) prior to harvesting roots for mRNA preparation, the potted trees were placed in a dark growth chamber at 25 C. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

#### ORIGIN

Query Match 5.2%; Score 39; DB 8; Length 742;  
Best Local Similarity 58.0%; Pred. No. 1.1;  
Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 398 TTAAGATGATGGGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 457  
DB 138 TTAAGATGAGGAT 197  
QY 458 ACATCACCATAAGAAATGACGATCCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 516  
DB 198 AACGAGGAT 256

RESULT 8  
LOCUS CO167205 782 bp mRNA linear EST 18-JUN-2004  
DEFINITION FLD1\_67\_C04.g1\_A029 Root flooded Pinus taeda cDNA clone  
LOCUS CO167205 FLD1\_67\_C04\_A029 5', mRNA sequence.  
ACCESSION CO167205  
VERSION CO167205.1 GI:48937746  
KEYWORDS EST.  
SOURCE Pinus taeda (loblolly pine)  
ORGANISM Pinus taeda

REFERENCE Buzaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 782)  
AUTHORS Pratt, L., Cordonnier-Pratt, M.M., Lorenz, W.W., Zimmermann, C., Johnson, H., Anfuoso, C., Kamran, D., Chhabra, D. and Dean, J.F.D.  
TITLE A loblolly pine (Pinus taeda) EST database from flooded roots unpublished (2004)  
JOURNAL Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

#### FEATURES

source Location/Qualifiers  
1..782  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="3 CCLONES"  
/db\_xref="taxon:3352"  
/clone="FLD1\_67\_C04\_A029"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Root flooded"  
/note="Organ: root; Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. prior to harvesting tissues for RNA isolation, the rooted cuttings were maintained for 27 days (April 2003) under



ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain 5% soil moisture content. Pots holding the rooted cuttings were fully submerged in water for 24 hours prior to harvest of the roots for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pBl1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

Query Match 5.2%; Score 39; DB 7; Length 782;

Best Local Similarity 58.0%; Pred. No. 1.1;

Matches 69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 398 TTAAAGTGTGGGATCGACGGGATAGTGAACCGCAATAGTGGCTTTGGATTATCA 457  
|||||  
DB 284 TTAAAGTGTGGGATCGTGGGAAGGAAGTTCTTTTGTGCTTTCAAGATGCA 343  
|||||

QY 458 ACATACCCATTAAGAAATGAACGATCCATCTCATATGAGTCCGTGATGGCATATTCAT 516  
|  
DB 344 AACAGGAAGGTGACATGTCTGCCAACATTAAAGAGGCCGTTAAGGCAAGTCTCCAT 402  
|

## RESULT 9

BU440426/c 371 bp mRNA linear EST 13-MAR-2002

LOCUS BU440426 Dictyostelium discoideum cDNA library, VF Dictyostelium  
DEFINITION dictoideum cDNA clone ddv43116 3', mRNA sequence.

ACCESSION BU440426  
VERSION BU440426  
KEYWORDS EST. GI:19415148

SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
Bukaryota; Mycetozoa; Dictyostellida; Dictyostelium.

REFERENCE 1 (bases 1 to 371)  
AUTHORS Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

TITLE Full length cDNA of Dictyostelium discoideum at the vegetative  
stage

JOURNAL Unpublished (2002)

COMMENT Contact: Tadaeu Shin-i  
Center For Genetic Resource Information

National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856  
Fax: 81-559-81-6855

Email: tshin@genes.nig.ac.jp.  
Location/Qualifiers

## FEATURES

source

1..371  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="ddv43116"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

## ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 371;

Best Local Similarity 51.4%; Pred. No. 1.2;

Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCGCAATAGTGTGGCTTATCAATCAACCATTAAGAAATGACGAT 481  
|||||  
DB 252 TGGTGAATAAAGAAATTAAGATTATGATCAACAAACCGATTAAGATGCAAT 193  
|||||

QY 482 CCACTCTATGAGAGTGTGATGATGATTCATTTGCAAACTTAAACGGAACTTT 541  
|||||  
DB 192 CTTTGTATATATATCAACCAAGGTGTTAGCTTACATTAAATGGGATATTGTTAAGATG 133  
|||||

QY 542 ATCAGCCGATTTCTACTGCTGAGATATATGCAAAATCAATTCAGGACTATT 594  
|

DB 132 CAACCTCAATGTTTACTGTGTGAAGTTTATGCTATTCATTAAGTCTGTAATAAT 80

## RESULT 10

BU440571/c

BU440571 Dictyostelium discoideum cDNA library, VF Dictyostelium

discoideum cDNA clone ddv44105 3', mRNA sequence.

ACCESSION BU440571 GI:19415293

VERSION

KEYWORDS

SOURCE

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

Dictyostelium discoideum

National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

## FEATURES

source

1..391  
/organism="Dictyostelium discoideum"  
/mol\_type="rRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="dv44m07"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

## ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 391;  
Best Local Similarity 51.4%; Pred. No. 1.2;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCCGAATGTCGTTGGTTATCACTCCATTAAGAAATGACGAT 481  
DB 248 TGGTTGAAAAAGAAATATAAAAGTTAATGATATCAACAAACCGTATTGATGCGAAT 189  
QY 482 CCACTCCATGAGAGTCGTTGATGCGATATCCATTGCGAAGCTTAAACGGAAAGTTT 541  
DB 188 CTTTGTTAATGATACCCAGGTGGTTTACGTTACCTTAATTAATGGGTATGTTAAAGATG 129  
QY 542 ATACACCGGATTCATCTGCTGATATATGCAATATCAAAATTCAGGACTATT 594  
DB 128 CAATTCAATGTTTACTGTTAGTAAGTTATGTCATTCATCAATGCTGTAAAT 76

RESULT 12 393 bp mRNA linear EST 13-MAR-2002  
LOCUS BJ440820/c Dictyostelium discoideum cDNA library, VF Dictyostelium  
DEFINITION BJ440820 Dictyostelium discoideum cDNA clone dv44g24 3', mRNA sequence.  
ACCESSION BJ440820.1 GI:19415542  
VERSION BJ440820.1  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.  
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.  
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage

Unpublished (2002)  
Contact: Tadao Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6855  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

## JOURNAL COMMENT

## FEATURES

source

1..393  
/organism="Dictyostelium discoideum"  
/mol\_type="rRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="dv44g24"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

## ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 393;  
Best Local Similarity 51.4%; Pred. No. 1.2;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCCGAATGTCGTTGGTTATCACTCCATTAAGAAATGACGAT 481  
DB 248 TGGTTGAAAAAGAAATATAAAAGTTAATGATATCAACAAACCGTATTGATGCGAAT 189  
QY 482 CCACTCCATGAGAGTCGTTGATGCGATATCCATTGCGAAGCTTAAACGGAAAGTTT 541  
DB 188 CTTTGTTAATGATACCCAGGTGGTTTACGTTACCTTAATTAATGGGTATGTTAAAGATG 129  
QY 542 ATACACCGGATTCATCTGCTGATATATGCAATATCAAAATTCAGGACTATT 594  
DB 128 CAATTCAATGTTTACTGTTAGTAAGTTATGTCATTCATCAATGCTGTAAAT 76

RESULT 13 394 bp mRNA linear EST 13-MAR-2002  
LOCUS BJ440720/c Dictyostelium discoideum cDNA library, VF Dictyostelium  
DEFINITION BJ440720 Dictyostelium discoideum cDNA clone dv44d15 3', mRNA sequence.  
ACCESSION BJ440720.1 GI:19415442  
VERSION BJ440720.1  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.  
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.  
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage

Unpublished (2002)  
Contact: Tadao Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.  
Location/Qualifiers  
1..394  
/organism="Dictyostelium discoideum"  
/mol\_type="rRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="dv44d15"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

## ORIGIN

Query Match 5.2%; Score 38.6; DB 3; Length 394;  
Best Local Similarity 51.4%; Pred. No. 1.2;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGGTGAACCCGAATGTCGTTGGTTATCACTCCATTAAGAAATGACGAT 481  
DB 253 TGGTTGAAAAAGAAATATAAAAGTTAATGATATCAACAAACCGTATTGATGCGAAT 194  
QY 482 CCACTCCATGAGAGTCGTTGATGCGATATCCATTGCGAAGCTTAAACGGAAAGTTT 541  
DB 193 CTTTGTTAATGATACCCAGGTGGTTTACGTTACCTTAATTAATGGGTATGTTAAAGATG 134  
QY 542 ATACACCGGATTCATCTGCTGATATATGCAATATCAAAATTCAGGACTATT 594  
DB 133 CAATTCAATGTTTACTGTTAGTAAGTTATGTCATTCATCAATGCTGTAAAT 81

RESULT 14 422 bp mRNA linear EST 13-MAR-2002  
LOCUS BJ437922/c Dictyostelium discoideum cDNA library, VF Dictyostelium  
DEFINITION BJ437922 Dictyostelium discoideum cDNA clone dv35e19 3', mRNA sequence.  
ACCESSION BJ437922.1 GI:19412644  
VERSION BJ437922.1  
KEYWORDS EST.

SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE 1 (bases 1 to 422)  
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.  
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadao Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
1..422  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="ddv35e19"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN  
Query Match 5.2%; Score 38.6; DB 3; Length 422;  
Best Local Similarity 51.4%; Pred. No. 1.2;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGTGAAACCGCAATGCGTTGGTTATCAACATCAAGTAAGAAATGACGAT 481  
DB 247 TGGTGAATAAAGAAATATAAAGTTATATATCAACAAACCGTATTAGATGCGAAT 188  
QY 482 CCATCTCTATGAGGCGTTGATGCAATTCATTGCAAGAACTTAAAGGGAAGCTTT 541  
DB 187 CTTTGTATTATATATCAACCAAGTGTTTACCTTACATTAATAATGGTATTGGTAAAGATG 128  
QY 542 ATCAGCCGAGTTCTACTGCTGAGATATATGCAAAATCAAAATTCAGGAGCTATT 594  
DB 127 CAATTCAATGTTTACTGTGAAGTTTATGCTCATTCAAATGCTGCTAAATAAT 75

RESULT 15  
BU444052/c 458 bp mRNA linear EST 13-MAR-2002  
LOCUS BU444052 Dictyostelium discoideum cDNA library, VF Dictyostelium  
DEFINITION dictoideum cDNA clone ddv55f06 3', mRNA sequence.  
ACCESSION BU444052  
VERSION BU444052.1 GI:19418773  
KEYWORDS EST.  
SOURCE Dictyostelium discoideum  
ORGANISM Dictyostelium discoideum  
REFERENCE 1 (bases 1 to 458)  
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.  
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative stage  
JOURNAL Unpublished (2002)  
COMMENT Contact: Tadao Shin-i  
Center for Genetic Resource Information  
National Institute of Genetics  
111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source  
1..458  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"

/clone="ddv55f06"  
/sex="mat A"  
/dev\_stage="Growth phase"  
/clone\_lib="Dictyostelium discoideum cDNA library, VF"

ORIGIN  
Query Match 5.2%; Score 38.6; DB 3; Length 458;  
Best Local Similarity 51.4%; Pred. No. 1.2;  
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 422 TAGTGAAACCGCAATGCGTTGGTTATCAACATCAAGTAAGAAATGACGAT 481  
DB 254 TGGTGAATAAAGAAATATAAAGTTATATATCAACAAACCGTATTAGATGCGAAT 195  
QY 482 CCATCTCTATGAGGCGTTGATGCAATTCATTGCAAGAACTTAAAGGGAAGCTTT 541  
DB 194 CTTTGTATTATATATCAACCAAGTGTTTACCTTACATTAATAATGGGTAATGTAAGAATG 135  
QY 542 ATCAGCCGAGTTCTACTGCTGAGATATATGCAAAATCAAAATTCAGGAGCTATT 594  
DB 134 CAATTCAATGTTTACTGTGAAGTTTATGCTCATTCAAATGCTGCTAAATAAT 82

Search completed: April 7, 2006, 22:37:13  
Job time : 3836 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 21:31:52 ; Search time 1537 Seconds

(without alignments)  
860.446 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744  
1 gtgcgaatcatgaagttt.....aaagacgtatgcttccg 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/1\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/H\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/PCUTS\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/ina/PP\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/ina/RE\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	33.8	4.5	601 3	US-09-949-016-150404 Sequence 150404, A
2	33.8	4.5	181251 3	US-09-949-016-15970 Sequence 15970, A
3	33.4	4.5	421118 3	US-09-949-016-16297 Sequence 16297, A
4	32.4	4.4	285478 3	US-09-949-016-13362 Sequence 13362, A
5	32.2	4.3	1141 3	US-09-806-708B-22 Sequence 22, Appl
6	32	4.3	832 3	US-09-621-976-2813 Sequence 2813, Ap
7	32	4.3	99304 3	US-09-949-016-15440 Sequence 15440, A
8	31.4	4.2	17000 3	US-09-679-299A-18 Sequence 18, Appl
9	31.4	4.2	236964 3	US-09-949-016-15753 Sequence 15753, A
10	31.2	4.2	8439 3	US-09-221-017B-473 Sequence 473, Appl
11	30.8	4.1	1237 2	US-08-808-793-26 Sequence 26, Appl
12	30.8	4.1	1995 3	US-09-710-279-1257 Sequence 1257, Ap
13	30.8	4.1	2169 3	US-09-134-001C-2645 Sequence 2645, Ap
14	30.8	4.1	2279 3	US-07-998-289B-3 Sequence 3, Appl1
15	30.8	4.1	3356 3	US-09-710-279-3920 Sequence 185, App
16	30.6	4.1	3001 3	US-09-539-333D-185 Sequence 185, App
17	30.4	4.1	601 3	US-09-949-016-151347 Sequence 151347, A
18	30.4	4.1	601 3	US-09-949-016-4254 Sequence 4254, Ap
19	30.4	4.1	3509 3	US-09-949-016-4255 Sequence 4255, Ap
20	30.4	4.1	3509 3	US-09-949-016-15996 Sequence 15996, A
21	30.4	4.1	118382 3	US-09-949-016-15997 Sequence 15997, A
22	30.2	4.1	1194 3	US-09-270-767-5598 Sequence 5598, Ap
23	30.2	4.1	1194 3	US-09-270-767-20880 Sequence 20880, A
24	30.2	4.1	1194 3	US-09-270-767-20880 Sequence 20880, A

C 25	30	4.0	601 3	US-09-949-016-128577 Sequence 128577, A
C 26	30	4.0	601 3	US-09-949-016-128578 Sequence 128578, A
C 27	30	4.0	601 3	US-09-949-016-150405 Sequence 150405, A
C 28	30	4.0	1270 3	US-09-415-277C-1 Sequence 1, Appl1
C 29	30	4.0	1270 3	US-09-415-277C-16 Sequence 16, Appl1
C 30	30	4.0	1271 3	US-09-415-277C-3 Sequence 3, Appl1
C 31	30	4.0	1522 3	US-09-533-559-5608 Sequence 5608, Ap
C 32	30	4.0	1641 2	US-09-080-897-5 Sequence 5, Appl1
C 33	30	4.0	1641 3	US-09-323-735-5 Sequence 5, Appl1
C 34	30	4.0	8019 3	US-09-949-016-2835 Sequence 2835, Ap
C 35	30	4.0	8019 3	US-09-949-016-2836 Sequence 2836, Ap
C 36	30	4.0	24205 3	US-09-949-016-15385 Sequence 15385, A
C 37	29.8	4.0	351 3	US-09-248-796A-2992 Sequence 2992, Ap
C 38	29.8	4.0	947 3	US-09-640-211A-138 Sequence 138, Appl
C 39	29.8	4.0	1164 3	US-09-248-796A-4125 Sequence 4125, Ap
C 40	29.8	4.0	1443 3	US-09-248-796A-668 Sequence 668, Appl
C 41	29.8	4.0	2090 3	US-09-270-767-4280 Sequence 4280, Ap
C 42	29.8	4.0	2090 3	US-09-270-767-19562 Sequence 19562, A
C 43	29.8	4.0	5619 3	US-09-799-451-241 Sequence 241, Appl
C 44	29.8	4.0	319608 3	US-09-539-333D-1 Sequence 1, Appl1
C 45	29.6	4.0	601 3	US-09-949-016-102623 Sequence 102623, A

# ALIGNMENTS

```

RESULT 1
US-09-949-016-150404
Sequence 150404, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150404
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-150404
Query Match
Beat Local Similarity 52.5%; Pred. No. 0.58; Length 601;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Cyt 186 TTTTGGTTAATTCACGAATAGTTCACAGTACGAAAGCAAACTCGAGTAC 245
Db 131 TTTTGGTTAATTCACGAATAGTTCACAGTACGAAAGCAAACTCGAGTAC 190
Cyt 246 GTTCGCGCTCATAGTACGAACTCTTCACATCGCAATCAAGTCTGACAC 305
Db 191 TTTTAAATTTTAAAGTACAGTACGAACTCTTCACATCGCAATCAAGTCTGACAC 250
Cyt 306 TTTTAAATTTTAAAGTACAGTACGAACTCTTCACATCGCAATCAAGTCTGACAC 326
Db 251 TTTTAAATTTTAAAGTACAGTACGAACTCTTCACATCGCAATCAAGTCTGACAC 271
RESULT 2
US-09-949-016-15970
Sequence 15970, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 15970
; LENGTH: 181251
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15970
```

Query Match 4.5%; Score 33.8; DB 3; Length 181251;

Best Local Similarity 52.5%; Pred. No. 13; Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```
Qy 186 TTCTTGGTATTGCGAATAGTTCACAGTAGCAAGCAAGCAAACTCGAGTAC 245
Db 58690 TTTTITTTTGAAGGCTGATTTTAAATGTAATAAGCATAAATTAATTTACTAT 58749
Qy 246 GTCTCCGCTCATAGTCTTACGAAGTCAATGCAACATCGCAATCAAGTGTCTACGAC 305
Db 58750 TTTTAAATTTTAAAGTGTACAGTACATTAACATTTACATCATGTGACGCTGTAC 58809
Qy 306 TTCATCTTCAGTCGGAAGT 326
Db 58810 TACCATTCATCTTGAAGCTT 58830
```

## RESULT 3

```
US-09-949-016-16297/C
; Sequence 16297, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 16297
; LENGTH: 42118
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)_(42118)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16297
```

Query Match 4.5%; Score 33.4; DB 3; Length 42118;

Best Local Similarity 49.2%; Pred. No. 28; Matches 88; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```
Qy 504 TGGCATATTCATTTGCAAGCACTTAACGGAACGTTTATGACCGGATTTACTGCTGA 563
Db 264214 TAGAATCATCATTTGTAAACAAATACGGCATGTTTCAACAAAGATTGACCTTTCA 264155
```

```
Qy 564 GATATGCAAAATCAATTCAGGACTATTCGACCGAAGCTCAACCTGGTGTGCAT 623
Db 264154 GGTATTTTCTATTATTCGCTAGATTAAGAAAGCTTAAGTATCTTGGGTATTA 264095
Qy 624 GACGACAAATCTCGTATTGATATGCGGATTTTATGACATATGTCAGTCGGGACG 682
Db 264094 CATTCACGAGAGGCTTTGAATAATTTTCCACACATGGAGTTATGTTAGC 264036
```

## RESULT 4

```
US-09-949-016-13362
; Sequence 13362, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13362
; LENGTH: 285478
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(285478)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13362
```

Query Match 4.4%; Score 32.4; DB 3; Length 285478;

Best Local Similarity 58.2%; Pred. No. 50; Matches 57; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy 249 TCCCGCTCATAGTCTGCAAGTCAATTCGAACATGCGCAATCAAGCTGTACGACTTC 308
Db 35275 TCTCCCTCTCAGGCTCTTAAGTATCTGAACGTGCGCACACACCCGGCTTAAGTTTT 35334
Qy 309 ATCTTTCAGTGAAGTTACATATATCCCGTTTCACA 346
Db 35335 GTATTTTCGTGAAGATGGGTTTCACCATGTTGCCCA 35372
```

## RESULT 5

```
US-09-806-708B-22/C
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURES:
; NAME/KEY: promoter
; LOCATION: (1)...(1141)
```



;; TITLE OF INVENTION: ANTISENSE MODULATION OF CASPASE 6 EXPRESSION  
;; FILE REFERENCE: RTS-0187  
;; CURRENT APPLICATION NUMBER: US/09/679,299A  
;; CURRENT FILING DATE: 2000-10-04  
;; NUMBER OF SEQ ID NOS: 164  
;; SEQ ID NO 18  
;; LENGTH: 17000  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-679-299A-18

Query Match 4.2%; Score 31.4; DB 3; Length 17000;  
Best Local Similarity 51.0%; Pred. No. 24;  
Matches 74; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

OY 441 TGGCTTTGGTTTCAACATCACCATAGAAATGAGCGATCCCTCTATGAGGTCT 500  
DB 16986 TGCCTTTCTTTTCCCTGACAGGCTAAGATCCCTGAAACAGCGCGCTCATTTCTTC 16927  
OY 501 TGATGCGATATTCATTTGCAAGAACTTAAACGGGACGTTATGACCGGATTTCTACTGC 560  
DB 16926 TGTTCGAATGCAAGTGAAGAACTCAATGAAAAGATTATCTTACAGTTTAAATGT 16867  
OY 561 TGAGATATATGCAAAATACAAATTC 585  
DB 16866 CGTCAGATTTTCATTTATGTGA 16842

RESULT 9  
US-09-949-016-15753/C  
Sequence 15753, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 15753

LENGTH: 236964

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)..(236964)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-15753

Query Match 4.2%; Score 31.4; DB 3; Length 236964;  
Best Local Similarity 50.3%; Pred. No. 98;  
Matches 77; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 451 GTTATCAATCATCACCATAAAGATGACGATCCCTCTATGAGGTCTGATGACATA 510  
DB 4225 GTTATTAATCTTGAATCATTAAGAAAAATTAAGAAATTAATGATGTCTATCTTCTATC 4166  
OY 511 TTCATTTCAGAACTTAAACGGGACGTTTATGACCGGATTTCTACTGCTGAGATATAT 570  
DB 4165 TTTCTCTACAGTTGAGAAACGCAATTTTCTAAGCTTATTAATTAAGAGTATAT 4106  
OY 571 GCAATTAAGAAATTCAGGAGCTATTTCCGACGAC 603  
DB 4105 AACCAATCTAGTTCACAAAAATATTTATGAGC 4073

RESULT 10  
US-09-221-017B-473

Sequence 473, Application US/09221017B

Patent No. 6444799

GENERAL INFORMATION:

APPLICANT: ROSES, BRUCE C.

TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF

NUMBER OF SEQUENCES: 1120

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 PAGE MILL ROAD

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows

SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/221,017B

FILING DATE: 23-DEC-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP1182

FILING DATE: 31-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP1546

FILING DATE: 30-JAN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP2911

FILING DATE: 09-APR-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU98/01023

FILING DATE: 10-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: MONROY, Gladys H.

REGISTRATION NUMBER: 32,430

REFERENCE/DOCKET NUMBER: 27340-20021.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-813-5600

TELEFAX: 650-494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 473:

SEQUENCE CHARACTERISTICS:

LENGTH: 8439 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: UNKNOWN

ORIGINAL SOURCE:

ORGANISM: PORPHYROMONAS GINGIVALIS

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1...8439

US-09-221-017B-473

Query Match 4.2%; Score 31.2; DB 3; Length 8439;  
Best Local Similarity 51.4%; Pred. No. 19;  
Matches 72; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

OY 378 AGTAACTCCCTGACAGAGTTAAGAGTGGGTACGACGGATAGGTGAACCGCAAA 437  
DB 5186 AGACAAATTCATTCAGGAATTCAGACGGCGCTTAATGGAACGGTGCAACGAAT 5245  
OY 438 TGTGCGCTTTGGTTTCAACATCACCATAAGAAATGACGAGATCCCTCTATGAGGT 497  
DB 5246 GGGATTGTGACAGTAATCAAAAAATATCTCGGCTCGGATGATCAATCTGGCAAAAAGC 5305



QY 498 CGTGAATGCAATTCATT 517  
DB 5306 CTATGCCGCAATTTGCTACT 5325

RESULT 11  
US-08-793-26/c  
Sequence 26, Application US/08808793  
Patent No. 5858713  
GENERAL INFORMATION:  
APPLICANT: Soderlund, David M.  
APPLICANT: Ingles, Patricia J.  
TITLE OF INVENTION: CALCIUM PERMEABLE INSECT SODIUM CHANNELS  
TITLE OF INVENTION: AND USE THEREOF  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,793  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/034,361  
FILING DATE: 24-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/012,649  
FILING DATE: 01-MAR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Braman, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/1062 (D-1906A)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1237 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-808-793-26

Query Match 4.1%; Score 30.8; DB 2; Length 1237;  
Best Local Similarity 52.3%; Pred. No. 9.2;  
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 3 GCGAATTATGAAAGTTTGGATGATTTGGTACGTAATATCATGCTTTATTCAGTG 62  
DB 862 GCGTAAAGCGCAATTTTAAATAGCACTTGAATCAATATCAATATATATCATGTT 803  
QY 63 GGGGCGAATGACAGCAATTTGCGTTTGGCCGTTTCATTTGGGCAATTAATTTGCTGC 122  
DB 802 GAGGATGTCAGGACAGTGTGAAGTCTCCGACTGTGTGTATGATGAGGCTCATGCT 743  
QY 123 GAACGCGTG 132  
DB 742 CAACATGTTG 733

RESULT 12

US-09-710-279-1257/c  
Sequence 1257, Application US/09710279  
Patent No. 6703492  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: PU3480US  
CURRENT APPLICATION NUMBER: US/09/710,279  
PRIOR FILING DATE: 2000-11-09  
PRIOR APPLICATION NUMBER: 60/164,258  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 1257  
LENGTH: 1995  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-710-279-1257

Query Match 4.1%; Score 30.8; DB 3; Length 1995;  
Best Local Similarity 52.3%; Pred. No. 12;  
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAATGCGCAATCAAGCTGCTACGACTTCATCTTCTGAGTCGAAGTTACGATA 332  
DB 263 ATTGCAATGCGCTGATTAATCTCAAGAACTAGTTTATTCATTCGTTTAAAGCG 204  
QY 333 TATCCGTTTACACATACCGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392  
DB 203 TATCCTTATAAATCTTCTGCGACATATATAGCAAAATGGAAGTAATGATTTT 144  
QY 393 GCAAGTTAG 402  
DB 143 GCGTATTGAG 134

RESULT 13  
US-09-134-001C-2645/c  
Sequence 2645, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIOR FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2645  
LENGTH: 2169  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2645

Query Match 4.1%; Score 30.8; DB 3; Length 2169;  
Best Local Similarity 52.3%; Pred. No. 12;  
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAATGCGCAATCAAGCTGCTACGACTTCAATCTTCTGAGTCGAAGTTACGATA 332  
DB 437 ATTGCAATGCTGATTAATCTCAAGAACTAGTTTATTCATTCGTTTAAAGCG 378  
QY 333 TATCCGTTTACACATACCGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392  
DB 377 TATCCTTATAAATCTTCTGCGACATATATAGCAAAATGGAAGTAATGATTTT 318

QY 393 GCAAGTTAAG 402  
|||  
Db 317 GCGATTTCAG 308  
^

## RESULT 14

US-07-998-289B-3/c  
; Sequence 3, Application US/07998289B  
; Patent No. 6027876  
; GENERAL INFORMATION:  
; APPLICANT: Black, Bruce C  
; APPLICANT: Taylor, Martin  
; APPLICANT: Heckel, David G  
; TITLE OF INVENTION: Method for Monitoring Pesticide  
; TITLE OF INVENTION: Resistance  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Darby & Darby PC  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: US  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/998,289B  
; FILING DATE: 30-DEC-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Robinson, Joseph R  
; REGISTRATION NUMBER: 33,448  
; REFERENCE/DOCKET NUMBER: 0646/OA939  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-527-7700  
; TELEFAX: 212-753-6237  
; TELEX: 236687  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2279 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-07-998-289B-3

Query Match 4.1%; Score 30.8; DB 3; Length 2279;  
Best Local Similarity 52.3%; Pred. No. 13;  
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 3 GCGAATTATGAAGTTTGGGATGATTTGGTACCGTATATCATGCTGTTTATTCAGTG 62  
|||  
Db 1462 GCGTAAAGGGAACATTTTAAATAGGACCTGAACGAAATATCAGATGAATATCATGTT 1403  
|||  
QY 63 GGGGCGAATGAACAGAAATTCGCGTTTGGCGGTTTCATTTGGGCAATTAATGTTCTTGGC 122  
|||  
Db 1402 GAGGTAGTCGAGACAGCTGCTGAAGTCTCCGACTGCTGTAAGTATGATGATGAGCGTATCGT 1343  
|||  
QY 123 GAAACGCGTG 132  
|||  
Db 1342 CAACATGTTG 1333  
|||

RESULT 15  
US-09-710-279-3920/c  
; Sequence 3920, Application US/09710279  
; Patent No. 6703452  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3920  
; LENGTH: 3356  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-3920

Query Match 4.1%; Score 30.8; DB 3; Length 3356;  
Best Local Similarity 52.3%; Pred. No. 16;  
Matches 68; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 273 ATTGCAACATGCGCAACATCAAGCTGCTAGACTTCATCTTCTGAGTGAAGTACGATA 332  
|||  
Db 666 ATTGCACTGCTGATTATTAATTCCTCAAGAACTTAGTTATCATTCGTTTAAAGCG 607  
|||  
QY 333 TATCCCGTTTCAACATACGGAAGTAGAGACTTGAATTAAGTTAACTCCCTGCA 392  
|||  
Db 606 TATCCTTATTAACCTTACTTGGCGACATATTACTAGCAAAATGGAAGTGAATGATTTT 547  
|||  
QY 393 GCAAGTTAAG 402  
|||  
Db 546 GCGATTTCAG 537  
|||

Search completed: April 8, 2006, 01:20:33  
Job time : 1540 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 22:13:32 ; Search time 823 Seconds  
(without alignments)  
7475.598 Million cell updates/sec

Title: US-10-784-592-18  
Perfect score: 744  
Sequence: 1 gtgcgaattacgaagttc.....aaacagcgtatccttcgcg 744

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	744	100.0	744	US-10-784-592-18	Sequence 18, Appl
2	36.2	4.9	389627	US-10-504-689-3	Sequence 3, Appl
3	36.2	4.9	405660	US-10-322-696-82	Sequence 82, Appl
4	35.2	4.7	580	US-10-487-901-4181	Sequence 4181, Ap
5	35	4.7	2349	US-10-369-963-37301	Sequence 37301, A
6	34.4	4.6	4818	US-10-437-963-63024	Sequence 63024, A
7	34.2	4.6	1502	US-09-925-065A-724617	Sequence 724617,
8	33.8	4.5	591	US-09-925-065A-867215	Sequence 867215,
9	33.8	4.5	660	US-10-425-115-125681	Sequence 125681,
10	33.8	4.5	729	US-10-425-114-6745	Sequence 6745, Ap
11	33.8	4.5	860	US-10-425-114-6834	Sequence 6834, Ap
12	33.8	4.5	935	US-10-425-115-154424	Sequence 154424,
13	33.8	4.5	1240	US-09-925-065A-41609	Sequence 41609, A
14	33.4	4.5	706	US-10-487-901-3702	Sequence 3702, Ap
15	33.4	4.5	1048	US-10-767-701-11172	Sequence 11172, A
16	33.4	4.5	12077	US-10-915-740A-281	Sequence 281, App
17	33.4	4.5	2242716	US-10-915-740A-1068	Sequence 1068, Ap
18	33.2	4.5	534	US-09-925-065A-427028	Sequence 427028,
19	33.2	4.5	640	US-09-925-065A-400132	Sequence 400132,
20	33	4.4	1184	US-10-123-155-412	Sequence 412, App
21	33	4.4	1184	US-10-146-731-412	Sequence 412, App
22	33	4.4	1184	US-10-140-472-412	Sequence 412, App
23	33	4.4	1184	US-10-141-761-412	Sequence 412, App

24	33	4.4	1184	US-10-142-885-412	Sequence 412, App
25	33	4.4	1184	US-10-158-790-412	Sequence 412, App
26	33	4.4	1184	US-10-137-871-412	Sequence 412, App
27	33	4.4	1184	US-10-140-923-412	Sequence 412, App
28	33	4.4	1184	US-10-141-756-412	Sequence 412, App
29	33	4.4	1184	US-10-141-759-412	Sequence 412, App
30	33	4.4	1184	US-10-140-805-412	Sequence 412, App
31	33	4.4	1184	US-10-140-864-412	Sequence 412, App
32	32.8	4.4	641	US-09-925-065A-178736	Sequence 178736,
33	32.6	4.4	566	US-09-925-065A-604703	Sequence 604703,
34	32.6	4.4	566	US-09-925-065A-604704	Sequence 604704,
35	32.6	4.4	2700	US-10-683-516-4	Sequence 4, Appl
36	32.4	4.4	611	US-09-925-065A-760902	Sequence 760902,
37	32.4	4.4	813	US-10-282-122A-16587	Sequence 16587, A
38	32.4	4.4	301477	US-10-322-281-456	Sequence 456, App
39	32.2	4.3	563	US-09-925-065A-192056	Sequence 192056,
40	32.2	4.3	598	US-09-925-065A-301137	Sequence 301137,
41	32.2	4.3	123526	US-09-910-185-11	Sequence 11, Appl
42	32	4.3	633	US-09-925-065A-914168	Sequence 914168,
43	32	4.3	1470	US-10-437-963-88649	Sequence 88649, A
44	32	4.3	26147	US-11-097-143-14227	Sequence 14227, A
45	31.8	4.3	503	US-10-363-345A-37281	Sequence 37281, A

ALIGNMENTS

```
RESULT 1
US-10-784-592-18
; Sequence 18, Application US/10784592
; Publication No. US20050147983A1
; GENBANK INFORMATION:
; APPLICANT: Willing, Reinhard
; APPLICANT: Ostergaard, Peter
; APPLICANT: Laessle, Soren
; TITLE OF INVENTION: POLYPEPTIDES OF ALICYCLOBACILLUS SP.
; FILE REFERENCE: 10406 203-US
; CURRENT APPLICATION NUMBER: US/10/784,592
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Alicyclobacillus sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(744)
; OTHER INFORMATION: s1g_peptide
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (124)..(744)
; OTHER INFORMATION: mat_peptide
US-10-784-592-18

Query Match      100.0%; Score 744; DB 9; Length 744;
Best Local Similarity 100.0%; Pred. No. 7.1e-240; Indels 0; Gaps 0;
Matches 744; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 GTGGGAATTATGAAGTTTGGATGATTTGGTACCGATATCATGCTTTATTCAG 60
Db      1 GTGGGAATTATGAAGTTTGGATGATTTGGTACCGATATCATGCTTTATTCAG 60
Cy      61 TTGGGGCGAATGAACAGAAATTCGCTTTTCCGCTTCATTTGGCGATTAAATGCTTC 120
Db      61 TTGGGGCGAATGAACAGAAATTCGCTTTTCCGCTTCATTTGGCGATTAAATGCTTC 120
Cy      121 GCGAACGCGGTATGATTCGAGAAACACCGCGGAACGATGAACGTAAGCGCT 180
Db      121 GCGAACGCGGTATGATTCGAGAAACACCGCGGAACGATGAACGTAAGCGCT 180
```



```

; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 69024
; LENGTH: 4818
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_69732C.1
US-10-437-963-69024

Query Match
Best Local Similarity 4.6%; Score 34.4; DB 7; Length 4818;
Matches 86; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Dh 558 TGCTGAATATATGCAATATACAAATTTCAGGAGCATATTCGACCGACTCAACCTGGTGT 617
939 TATCGAATCTAGCTAGAAAGGAGAAAGGGGGGGAATATTTAACTGTACCCAGATCTTGATGT 998

Dy 618 GTCCATACACCAAAATCTCGTATTTGATATGCCGATTTTATGACATATGGTCACTCGG 677
999 TTACATAGAGCAATATCACTGGTGGTGTCTAGAGACAAATATATCTACTGATTTACGTCTT 1058

Dy 678 GCAGCATTAATCACTTGTGCTTCATAGGGTTTCTTGCGGTCAAGTAAACG 729
Dh 1059 CAAGATTCTAGGACTTGACATTTGCGCTGACACCATTTGTTGAAATGAAATG 1110

RESULT 7
US-09-925-065A-724617/C
; Sequence 724617, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OR INVENTION: Identification and Mapping of Single
; TITLE OR INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: PaSeq for Windows Version 4.0
; SEQ ID NO 724617
; LENGTH: 1502
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-724617

Query Match
Best Local Similarity 4.6%; Score 34.2; DB 4; Length 1502;
Matches 120; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

Dh 320 CGAAGTTACGATATATTCCTTCCTTCACACATACCGGAAGATATGAGACTTGAAATTAAG 379
783 CTATGTAGAGCATTTCTTGCGACTGCTCAAGACCGAAGAAAGGCTGTGTGATGAG 724

Dy 380 TTAATCTCCCTGACGAAGTTAAGATGTGGGTTACAGACCGGATAGTGAACCGCAATG 439
723 TGAAGTGAAGGGGTGTGGAAGAAAGTATATATGTGTGAAGATGTAAACAGAGGTC 664

Dy 440 GTGCGTTTGGGTATACATCACTAATTAAGAAATGACGATCACTCTTATGAGGTG 499
Dh 663 AGATTAATTAAGGGCTTGAAGGCCACATATGAATTTTGGCTTTTCTTCAAGTTATATAG 604

```

Oy 500 TTGAATGGCATATTCATTTTCAGAACTTAAAGGGAACTTTATTCAGCCGAGATTCTACTG 559  
 Db 603 AAGAAGAAATATTCAAATTTTAAAACTTTCTCTCATGAATTTAAATAAAGAAATACATGA 544  
 Oy 560 CTGAGATATATGCAAAATACAAAT 582  
 Db 543 AAATGAAATGTACATCTCATATAT 521

## RESULT 8

```

US-09-925-065A-867215
? Sequence 867215, Application US/09925065A
? Publication No. US20050228172A9
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single
? FILE REFERENCE: 108827.135
? CURRENT APPLICATION NUMBER: US/09/925,065A
? PRIOR FILING DATE: 2001-08-08
? PRIOR APPLICATION NUMBER: US 60/243,096
? PRIOR FILING DATE: 2000-10-24
? PRIOR APPLICATION NUMBER: US 60/252,147
? PRIOR FILING DATE: 2000-11-20
? PRIOR APPLICATION NUMBER: US 60/250,092
? PRIOR FILING DATE: 2000-11-30
? PRIOR APPLICATION NUMBER: US 60/261,766
? PRIOR FILING DATE: 2001-01-16
? PRIOR APPLICATION NUMBER: US 60/289,846
? PRIOR FILING DATE: 2001-05-09
? NUMBER OF SEQ ID NOS: 957086
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 867215
? LENGTH: 591
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-925-065A-867215

```

Query Match	4.5%;	Score 33.8;	DB 4;	Length 591;
Best Local Similarity	52.5%;	Pred. No. 4.8;		
Matches	74;	Conservative	0;	Mismatches 67;
				Indels 0;
				Gaps 0

Oy	186	TTCTTTTGTTAATTGACGAATAGTTCACAGGTAGCAATGCCAAGACAAACTCGAGTAC	245
Db	289	TTTTTTTTTGAAGTGGGCTCATTTTTTTAAATTGTAAATTAAGCATTAATAAATTTTCTAT	348
Oy	246	GTCCTCCCGCTATAAGTCTACGAACCTCATTTGCACAATGGGCAACATCAAGTGTCTACGAC	305
Db	349	TTTAAACAATTTTAAAGTAGTACAGTAGACATTACACAACATTTACATCAATGTGCACGTCTTAC	408
Oy	306	TTTCATCTTCTCGAGTGAAGTT	326
Db	409	TACCAATCAATCTCTAGAACTT	429

## RESULT 9

```

US-10-425-115-125681
: Sequence 125681, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B
: CURRENT APPLICATION NUMBER: US/10/425,115
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 369326
: SEQ ID NO 125681
:
: LENGTH: 660

```

```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_46091C.1
US-10-425-115-125681

```

Query Match	4.5%	Score 33.8;	DB 8;	length 660;
Best Local Similarity	52.5%;	Pred. No. 5.1;		
Matches 74;	Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0

QY	584	CAGGAGCATTTTCCGACCGACCTCAACCCCTGGAGTGTCTCATGACGACAAATTCCTGTTATTG	643
Db	391	CTGGGCATATCTGCACCGGAGTGTCAAGGCCACGGAAGTCCAAACGATCTTTAGGAAATGGG	450
QY	644	AATATGCCGATTTTATATGACATATGGTACACGTGGGCGACGATTAATCTATCTGTCGCTTCCA	703
Db	451	TTTACCGGGAATCTCTCTGATATATACGTGTCTGAGAGAAATTAAGTACGTGTGCTCTCTTA	510
QY	704	TGGGTTTCTTCGGGTCAATG	724
Db	511	TGAGTGTAGGCCGAGCACTACG	531

## RESULT 10

```

US-10-425-114-6745
; Sequence 6745, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6745
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700577561_FLI
US-10-425-114-6745

```

Query Match	4.5%	Score 33.8;	DB 7;	Length 729;
Best Local Similarity	52.5%;	Pred. No. 5.4;		
Matches 74;	Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0;

OY	584	CAGGACATTTTCCACCGACCTTCAACCTTGCTGTGTCCATGACGACAAATCTCGATTGTTG	643
Db	464	CTGGGCATATCTGCACCGGAGTACAAAGCCCAAGAGTCCAAACGATTTAGGAAATGAG	523
OY	644	ATATCCCGAATTTTATGACATATGTCACGTGGGCGACGATTACTCATTTGTGCTTCCA	703
Db	524	TTTCAACGGGATCTCTCTGATATTACGTGCTGAGAGATTAACGTACGTGCTCTCTA	583
OY	704	TGGGTTTTCTTGGGTCAGATG	724
Db	584	TGAGTGTAGGCGCGACGTACG	604

## RESULT 11

US-10-425-114-6834  
; Sequence 6834, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yima  
; APPLICANT: Kovacic, David K.

```

; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 6834
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700579892_FLI
US-10-425-114-6834

Query Match
Best Local Similarity 4.5%; Score 33.8; DB 7; Length 860;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTCCGACCGACCTCAACCTGTGTGTCCATGACGACCAATCTCGTATTG 643
DB 600 CTGGGCATATCGCACCGGAGTCAAGCCACGAGAGTGCACAAAGATGTTAGAAATGG 659
QY 644 ATATGCCGAGTTTATGACATATGATCACTGCGGACAGCATTCTCACTTGTGCTTCA 703
DB 660 TTTCACGGGATCTCTCTGATATATGCTGTCTGAGAGAAATTCGTACGTGCTCTCA 719
QY 704 TGGGTTCTTGGGTCAGATG 724
DB 720 TGAAGTAGCCCGACGTAACG 740

RESULT 12
US-10-425-115-154424
; Sequence 154424, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 154424
; LENGTH: 935
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MNT4577_72415C.1
US-10-425-115-154424

Query Match
Best Local Similarity 4.5%; Score 33.8; DB 8; Length 935;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 584 CAGGACTATTCCGACCGACCTCAACCTGTGTGTCCATGACGACCAATCTCGTATTG 643
DB 593 CTGGGCATATCGCACCGGAGTCAAGCCACGAGAGTGCACAAAGATGTTAGAAATGG 652
QY 644 ATATGCCGAGTTTATGACATATGATCACTGCGGACAGCATTCTCACTTGTGCTTCA 703
DB 653 TTTCACGGGATCTCTCTGATATATGCTGTCTGAGAGAAATTCGTACGTGCTCTCA 712
QY 704 TGGGTTCTTGGGTCAGATG 724
DB 713 TGAAGTAGCCCGACGTAACG 733
```

```

RESULT 13
US-09-925-065A-41609
; Sequence 41609, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41609
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-41609
```

```

Query Match
Best Local Similarity 4.5%; Score 33.8; DB 4; Length 1240;
Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```

```

QY 484 ACTCCATGAGGTCGTGATGAGATTCATTTGACAACTTGAAGGAGAGCTTAT 543
DB 145 ACTCTTACTACAGATGCTACTGTTCTGATGACCAAGATTAATACTGAATTAA 204
QY 544 CAGCCGAGTTTACTGTCGATATATGCAATATCAAAATTCAGGACTATTCGACGAC 603
DB 205 AATCTCTTCACTGTCGTATATATATATATATATATATATATATATATATATAT 264
QY 604 CTCAACCTGTGTGTCATATGACGACAAATCTCGTAT 640
DB 265 GTCAACATTCAGTGAAGAAAGATGTCATCAGGAT 301
```

```

RESULT 14
US-10-487-901-3702/c
; Sequence 3702, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCrery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thadeus
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinna, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Crosley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Characteri
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3702
; LENGTH: 706
; TYPE: DNA
; ORGANISM: Artificial
```





November 2005

Published\_Applications\_Nucleic Acid and Published\_Applications\_Amino Acid database searches now generate two sets of results each. The Published\_Applications\_databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases: older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).  
Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (uspto)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 22:17:35 ; Search time 531 Seconds  
(without alignments)

5605.299 Million cell updates/sec

Title: US-10-784-592-18

Perfect score: 744

Sequence: 1 gtgcgaattatgaagtttt.....aaacgacgatgtctcttcgcg 744

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 200278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /SIDSS7/ptodata/1/pubpna/US08 NEW PUB.seq.\*
- 2: /SIDSS7/ptodata/1/pubpna/US06 NEW PUB.seq.\*
- 3: /SIDSS7/ptodata/1/pubpna/US07 NEW PUB.seq.\*
- 4: /SIDSS7/ptodata/1/pubpna/PCT NEW PUB.seq.\*
- 5: /SIDSS7/ptodata/1/pubpna/US09 NEW PUB.seq.\*
- 6: /SIDSS7/ptodata/1/pubpna/US09 NEW PUB.seq.\*
- 7: /SIDSS7/ptodata/1/pubpna/US10 NEW PUB.seq.\*
- 8: /SIDSS7/ptodata/1/pubpna/US10 NEW PUB.seq.\*
- 9: /SIDSS7/ptodata/1/pubpna/US10 NEW PUB.seq.\*
- 10: /SIDSS7/ptodata/1/pubpna/US10 NEW PUB.seq.\*
- 11: /SIDSS7/ptodata/1/pubpna/US11 NEW PUB.seq.\*
- 12: /SIDSS7/ptodata/1/pubpna/US11 NEW PUB.seq.\*
- 13: /SIDSS7/ptodata/1/pubpna/US11 NEW PUB.seq.\*
- 14: /SIDSS7/ptodata/1/pubpna/US11 NEW PUB.seq.\*
- 15: /SIDSS7/ptodata/1/pubpna/US60 NEW PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
c 1	34.2	4.6	1502	6 US-09-925-065A-724617
c 2	33.8	4.5	591	6 US-09-925-065A-867215
c 3	33.8	4.5	1240	6 US-09-925-065A-41609
c 4	33.8	4.5	1240	9 US-10-301-480-142847
c 5	33.8	4.5	1240	8 US-10-301-480-756256
c 6	33.4	4.5	1848	8 US-10-467-657-1219
c 7	33.2	4.5	534	6 US-09-925-065A-427028
c 8	33.2	4.5	534	10 US-10-301-480-490436
c 9	33.2	4.5	534	10 US-10-301-480-1103845
c 10	33.2	4.5	640	6 US-09-925-065A-400132
c 11	33.2	4.5	643	10 US-10-301-480-467470
c 12	33.2	4.5	643	10 US-10-301-480-1080879
c 13	32.8	4.4	641	6 US-09-925-065A-178736
c 14	32.8	4.4	650	10 US-10-301-480-269706
c 15	32.8	4.4	650	10 US-10-301-480-883115
c 16	32.6	4.4	566	6 US-09-925-065A-604703
c 17	32.6	4.4	566	6 US-09-925-065A-604704
c 18	32.4	4.4	611	6 US-09-925-065A-760902

c 19	32.2	4.3	563	6 US-09-925-065A-192056
c 20	32.2	4.3	571	10 US-10-301-480-281632
c 21	32.2	4.3	571	10 US-10-301-480-895041
c 22	32.2	4.3	598	6 US-09-925-065A-301137
c 23	32.2	4.3	633	6 US-09-925-065A-914168
c 24	31.8	4.3	528	10 US-10-301-480-430289
c 25	31.8	4.3	528	10 US-10-301-480-1043698
c 26	31.8	4.3	533	6 US-09-925-065A-413580
c 27	31.8	4.3	551	10 US-10-301-480-479582
c 28	31.8	4.3	551	10 US-10-301-480-1092991
c 29	31.8	4.3	557	10 US-10-301-480-492045
c 30	31.8	4.3	557	10 US-10-301-480-1105454
c 31	31.8	4.3	558	6 US-09-925-065A-359103
c 32	31.8	4.3	573	6 US-09-925-065A-429103
c 33	31.8	4.3	615	6 US-09-925-065A-517010
c 34	31.8	4.3	1400	14 US-11-136-527-6130
c 35	31.8	4.3	4953	14 US-11-136-527-2034
c 36	31.6	4.2	633	14 US-11-000-463-104
c 37	31.6	4.2	634	14 US-11-000-463-576
c 38	31.6	4.2	1394	6 US-09-925-065A-711186
c 39	31.4	4.2	540	6 US-09-925-065A-747461
c 40	31.4	4.2	545	6 US-09-925-065A-746148
c 41	31.4	4.2	611	6 US-09-925-065A-226317
c 42	31.2	4.2	505	10 US-10-301-480-471976
c 43	31.2	4.2	505	10 US-10-301-480-1085385
c 44	31.2	4.2	524	6 US-09-925-065A-405101
c 45	31.2	4.2	1173	8 US-10-517-939-27

ALIGNMENTS

RESULT 1

US-09-925-065A-724617/c  
; Sequence 724617, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925.065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 724617  
; LENGTH: 1502  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-724617

Qy	320	CGAAGTTACGATATATCCGTTTCACATACCGGAGGTAGGACCTTGGAATTAGAG	379
Dy	783	CTAATGTAGGAGCATTTCTGCGATGCTCAAAGACCAGCAAGAGCCTGTGTGATAGAG	724
Qy	380	TTAATCTCTGCGAGCATTTAAGAGTGTGGGTAGCGGNTAGGTGAACCCGAATG	439
Dy	723	TGAAGTGGATGGGTGTGTTGAAAGGAAGTAGTATATATGTTGTAAGAGATGTATACAGGAGGTC	664

Qy	440	GTGCGTTTGGGTATCAACATCACCATAGAAATGACGATCCACTCTATGGAAGTCTG	499
Db	663	AGATAATTAGGCTTTGAAGCCACCAATAAGAAATTTTGGCTTTTCTTCAAGTTATATAG	604
Qy	500	TTGATGGCATATTCGATTTGGCAGAACTTAAACGGGAACGTTTATCAGCGGAGTTCTCTG	559
Db	603	AAGAAGAATAATTCAAATTTTTTAAACTTTTCTGTCATGAAATTATAATAAGAATACATGA	544
Qy	560	CTGAGATATATGCAATACAAAT	582
Db	543	AAATGAAATGTACATCTCAATAT	521

## RESULT 2

```

US-09-925-065A-867215
; Sequence 867215, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Polymorphic Regions in the Human Genome
; TITLE OF INVENTION: Nucleotide Polymorphism in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 867215
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-867215

```

	Query Match	4.5%;	Score 33.8;	DB 6;	Length 591;
	Best Local Similarity	52.5%;	Pred. No. 3.1;		
	Matches 74;	Conservative 0;	Mismatches 67;	Indels 0;	Gaps 0;
Qy	186	TTCTTTGGTTAAATTCGACGAAATAGTTACACAGGTAGCAAGCGACGAAACCTCGAGTAC	245		
Db	289	TTTTTTTTTTCGAGTGGGTCTATTTTAAAAATTGTAATAAGCATCAACATAAAATTTACTAT	348		
Qy	246	GTCTCCCGCTCATAAAGTCTAGCAACTCATTTGCAACATCGCGCAACATCAAGCTGCTACGAC	305		
Db	349	TTTAAACAATTTTAAAGTGTACAGTAGCATTAACAACATTTATCATCTTGTGCGAGCTGTAC	408		
Qy	306	TTCATCTTCTCAGTCGAAGTT	326		
Db	409	TACCATGCATCTCTAGAAGCTT	429		

### RESULT 3

```

US-09-925-065A-41609
; Sequence 41609, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147

```

```

; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41609
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-41609

```

	Query Match	4.5%	Score 33.8	DB 6	Length 1240
	Best Local Similarity	51.0%	Pred. No. 4.4		
	Matches	80	Conservative	0	Mismatches 77; Indels 0; Gaps 0;
Qy	484	ACTCCTATGAGGTCGTGTGATGGCATATTC	CAATTCGAGAACTTAAACGGGACGTTTAT	543	
Db	145	ACTCTTTTACTCAGCATGCTACTGTTCTTGATGACCAAGATAAAAAC	TGAACCTTTTAAA	204	
Qy	544	CAGCGGATCTCTCTGAGATATATGCAAAATACAAATTCAGGAGCTATTCCGACCGCAC	603		
Db	205	AATCTCTCTCCTGTGTGTATATATATAAAAAAATACTGGACCTTTATCACCACA	264		
Qy	604	CTCAACCTGGTGTGTCATGACGACAAATCTCGTAT	640		
Db	265	GTCAACAATCCAGTGAAGAAGATGGTCACATCAGGAAT	301		

## RESULT 4

```

US-10-301-480-142847
; Sequence 142847, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108927.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 142847
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-142847

```

	Query Match	4.5%	Score 33.8	DB 9	Length 1240
	Best Local Similarity	51.0%	Pred. No. 4.4		
	Matches 80	Conservative 0	Mismatches 77	Indels 0	Gaps 0
Qy	484	ACTCCTATGAGGTCTGTGTGGCATATTCATTTGCAGAACTTAAACGGGAAGTTTAT	543		
Db	145	ACTCTTTTACTCAGCATGCTACTGTTTCTGTATGACCAAAAGATAAAACCTTTTAAA	204		
Qy	544	CAGCGGATTTCTACTGCTGAGATATATGCAAAATACAAATTCAGGGACTATTCCGACCGAC	603		
Db	205	AATCTCTTTCACCTGTGTGTATATATATAAAAAAATACTGGACCTTTATCACCATA	264		
Qy	604	CTCAACCTCGTGTGTCATGACGACAAATCTCGTAT	640		
Db	265	GTCAACAATCCAGTGAAGAAGATGGTCACATCAGGAAT	301		

## RESULT 5

```
US-10-301-480-756256
; Sequence 756256, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 756256
; LENGTH: 1240
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-756256

Query Match          4.5%; Score 33.8; DB 10; Length 1240;
Best Local Similarity 51.0%; Pred. No. 4.4;
Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

Qy 484 ACTCGATGAGGTCGTTGATGGCATATTCATTTGCAGAACTTAAACGGGAACGTTTAT 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 ACTCTTTTACCTCAGATGCTACTGTTCTGATGACCAAGATAAATACTGAACCTTTAA 204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 544 CAGCGGATTTACTGTGAGATATATGCAATATCAAAATTCAGGACTATTCCGACCGAC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 AATCTCTCTTCCACTGTGTATATATATAAAAAAATACTGACCTTTTATCACCAAA 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 604 CTCACCTGTGTGTCTCATGACGACAAATCTCGTAT 640
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 GTCAACAATCCAGTGAAGAAAGATGGTCACATCAGGAAT 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-10-467-657-1219
; Sequence 1219, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1219
; LENGTH: 1848
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1219

Query Match          4.5%; Score 33.4; DB 8; Length 1848;
Best Local Similarity 65.3%; Pred. No. 7.1;
Matches 49; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 122 CGAACACGGGTATATATGATTCGAGAAACACACCGCGGAACGCATCAACGGTAAGCGTAA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1281 CGACACGGTTTTTACGATATGGGGTCAACAGCGCGGAACCGACCGCAACTCAAA 1340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 182 CAACTTCTTTGGTTA 196
    ||||| ||||| |||||
```

```
Db 1341 CTACACCTTGGGCTA 1355

RESULT 7
US-09-925-065A-427028/c
; Sequence 427028, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427028
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-427028

Query Match          4.5%; Score 33.2; DB 6; Length 534;
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTTAAGAGTGTGGGGTACGCGGATAGGTGAACCGCA 435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 ACATTTAACAGGCTGCAAGAAGTTTGCCTGCTCTCCCACTGGGTAGCGTCAGGACT 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 436 AATGGTGGTTTTTGGGTTATCAATCACCATAAGAATGACGGATCCATCTCTATGGAG 495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 CACAGAGAGTCAGGCGCTTTTCATCATTTACACAGAGTAAGAGCCATTCACCCCATGTG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 496 GTCGTTGATGCGCATAT 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 CCTGTGTGGCCATGT 56
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-301-480-490436/c
; Sequence 490436, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 490436
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-490436

Query Match          4.5%; Score 33.2; DB 10; Length 534;
```

```
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTAAGAGTGTGGGTACGACGGGATAGGTGAACCGCA 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ACATTTAACAGGCTGCAAGAAGTTGCCTGCTCTCCCAATGGGGTAGCGTCAGGACT 132
Qy 436 AATGGTGGCTTTGGGTTATCAATCACCATAAGAAATGACGGATCCACTCCTATGGAG 495
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 CACAGAGAGTCAGGCTTTTTCATCATTAACAGAGTAAGAGCCATTACCCCCCATGGT 72
Qy 496 GTCGTGATGGCATAT 511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 CCTGTGTGGCCATGT 56

RESULT 9
US-10-301-480-1103845/c
; Sequence 1103845, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1103845
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1103845

Query Match 4.5%; Score 33.2; DB 10; Length 534;
Best Local Similarity 52.2%; Pred. No. 4.6;
Matches 71; Conservative 1; Mismatches 64; Indels 0; Gaps 0;

Qy 376 AGAGTTAACTCCCTGCAGCAAGTTAAGAGTGTGGGTACGACGGGATAGGTGAACCGCA 435
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 ACATTTAACAGGCTGCAAGAAGTTGCCTGCTCTCCCAATGGGGTAGCGTCAGGACT 132
Qy 436 AATGGTGGCTTTGGGTTATCAATCACCATAAGAAATGACGGATCCACTCCTATGGAG 495
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
131 CACAGAGAGTCAGGCTTTTTCATCATTAACAGAGTAAGAGCCATTACCCCCCATGGT 72
Qy 496 GTCGTGATGGCATAT 511
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
71 CCTGTGTGGCCATGT 56

RESULT 10
US-09-925-065A-400132/c
; Sequence 400132, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
```

```
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 400132
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-400132

Query Match 4.5%; Score 33.2; DB 6; Length 640;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TCGCAATTATGAAAGTTTGGGATGGATTTTGGTACCGTATATCATGCTGTTTATTCAGT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
313 TTCTTATATTCAAAGTATTGACTTGTATCTTTGTACCTTATATTGATGTTTATTCAGT 254
Qy 62 GGGGGCGAATGAACAGAAATT 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GTGTTTCAAATCAAAGTTT 234

RESULT 11
US-10-301-480-467470/c
; Sequence 467470, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467470
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-467470

Query Match 4.5%; Score 33.2; DB 10; Length 643;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TCGCAATTATGAAAGTTTGGGATGGATTTTGGTACCGTATATCATGCTGTTTATTCAGT 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
313 TTCTTATATTCAAAGTATTGACTTGTATCTTTGTACCTTATATTGATGTTTATTCAGT 254
Qy 62 GGGGGCGAATGAACAGAAATT 81
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
253 GTGTTTCAAATCAAAGTTT 234

RESULT 12
US-10-301-480-1080879/c
; Sequence 1080879, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 60/311,695
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1080879
; LENGTH: 643
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1080879
```

```
Query Match          4.4%; Score 33.2; DB 10; Length 643;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 50; Conservative 1; Mismatches 29; Indels 0; Gaps 0;

Qy 2 TCGCAATTATGAAGCTTTTGGATGCAATTTTGGTACCGTATATCATGCTGTTTATTTCAGT 61
Db 313 TTCTATATTCAAGTATGACTTGTACTTTGTACCTTATATTGATGTTTATTCACT 254

Qy 62 GGGGCGGAATGAACGAATT 81
Db 253 GTGTTTCAATCAAGTTTY 234
```

```
RESULT 13
US-09-925-065A-178736/c
; Sequence 178736, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 178736.1
; LENGTH: 641
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-178736
```

```
Query Match          4.4%; Score 32.8; DB 6; Length 641;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTACGACTTCATCTTCTCAGTCGAAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTCTTCATTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCGACGAAGTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAACTTACATAAATTTCCAGCATTCACTTTAAGCTGTCCGATAC 503
```

```
RESULT 14
US-10-301-480-269706/c
; Sequence 269706, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
```

```
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 269706
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-269706
```

```
Query Match          4.4%; Score 32.8; DB 10; Length 650;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTACGACTTCATCTTCTCAGTCGAAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTCTTCATTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCGACGAAGTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAACTTACATAAATTTCCAGCATTCACTTTAAGCTGTCCGATAC 503
```

```
RESULT 15
US-10-301-480-883115/c
; Sequence 883115, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 883115
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-883115
```

```
Query Match          4.4%; Score 32.8; DB 10; Length 650;
Best Local Similarity 55.2%; Pred. No. 6.8;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 299 CTACGACTTCATCTTCTCAGTCGAAGTTACGATATATCCCGTTTTCACACATACGGGAAGG 358
Db 618 CTTTGCTCTTCATTTTAAATATAAATACTAGGATATCACTTGTTCCTTTGCTTCTTGG 559

Qy 359 TAGGAGACTTGGAAATTAGAGTTAACTCCCTGCGACGAAGTTAAGAGTGTGGGTAC 414
Db 558 TATAGCGCTGTAACTTACATAAATTTCCAGCATTCACTTTAAGCTGTCCGATAC 503
```

Search completed: April 7, 2006, 22:41:16  
Job time : 533 secs

This page blank (uspro)





CC 15716. The invention is useful in industrial, research and household  
CC processes such as in composition to prepare feed and food, in detergent  
CC formulations and for treating lignocellulosic fabric and pulp. The  
CC present sequence is Alicyclobacillus sp. DSM 15716 functional  
CC polypeptide.

XX SQ Sequence 248 AA;

Query Match 100.0%; Score 1289; DB 9; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3e-108;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVFANTYVMIRGTPRNASTVSA 60  
DB 1 KRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVFANTYVMIRGTPRNASTVSA 60  
QY 61 TTSLVNSTNSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSKLRYPFTYGV 120  
DB 61 TTSLVNSTNSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSKLRYPFTYGV 120  
QY 121 GDLEIRVNSLOQVKSVDGIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180  
DB 121 GDLEIRVNSLOQVKSVDGIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180  
QY 181 YQDSTAEIYANTNSGTIPDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240  
DB 181 YQDSTAEIYANTNSGTIPDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240  
QY 241 DDTYALP 248  
DB 241 DDTYALP 248

#### RESULT 2

ID AEB48755  
AEB48755 standard; protein; 248 AA.

AC AEB48755;  
XX  
DT 22-SEP-2005 (first entry)

DE Alicyclobacillus sp. DSM 15716 functional polypeptide.

KM Protein production; protein secretion; surfactant; feedstuff; food.

OS Alicyclobacillus sp.; 'DSM 15716'.

FX Key Location/Qualifiers

FT Peptide 1..41

FT Misc-difference 1 /label= Signal\_peptide

FT Protein 42..248 /note= "Encoded by GTC"

FT Protein /label= Mature\_polypeptide

XX WO200506339-A2.

XX 21-JUL-2005.

XX 06-JAN-2005; 2005WO-DK000004.

XX 06-JAN-2004; 2004DK-00000010.

XX 04-FEB-2004; 2004DK-00000165.

XX 23-FEB-2004; 2004US-00784592.

XX 25-FEB-2004; 2004DK-00000298.

XX (NOVO) NOVOZYMES AS.

XX Wilting R, Laasen SF, Ostergaard PR;

XX MPI; 2005-506869/51.

XX N-PSDB; AEB48730.

PT New mature functional polypeptide of Alicyclobacillus sp., used for  
PT commercial research purposes, specifically for cleaning a cellulosic  
PT fabric, preparing food or feed additive, and for treating lignocellu-  
PT materials and pulp.

PS Claim 3; SEQ ID NO 43; 151pp; English.

XX The present invention is based on the finding of a strain of  
CC Alicyclobacillus, namely Alicyclobacillus sp. DSM 15716, which grows at  
CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It  
CC is an object of the invention to identify and provide polypeptides  
CC secreted from Alicyclobacillus sp. DSM 15716 because such polypeptides  
CC may be used for industrial purposes and may also be produced in  
CC industrially relevant processes and amounts. Thus, the invention provides  
CC isolated mature functional polypeptides which are at least 90% identical  
CC to, and exhibit and same function as a corresponding secreted polypeptide  
CC AEB48738-AEB48762 from Alicyclobacillus sp. DSM 15716, and also provides  
CC polynucleotides AEB48733-AEB48737 encoding such polypeptides. The  
CC polypeptides are preferably enzymes having acid endoglucanase, acid  
CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl  
CC proase, serine protease, HtrA-like serine protease, diulfide  
CC isomerase, gamma-D-glutamyl-L- amino acid endopeptidase, endo-beta-N-  
CC acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase,  
CC phytase, phospholipase C, polysaccharide deacetylase, xylanase  
CC deacetylase or sulfite oxidase activity. A composition comprising such an  
CC enzyme can be used in a detergent composition, or a food or feed  
CC composition. Also claimed are: nucleic acid constructs, recombinant  
CC expression vectors and host cells; methods of producing the polypeptides;  
CC a storage medium suitable for use in an electronic device comprising  
CC information of the amino acid sequence of a polypeptide of the invention  
CC or the nucleotide sequence encoding the polypeptide; and a process  
CC comprising employing the polypeptide, or polynucleotide encoding it, in  
CC an industrial or household technical process. The present sequence is the  
CC protein sequence of a functional polypeptide of the invention.

XX SQ Sequence 248 AA;

Query Match 100.0%; Score 1289; DB 9; Length 248;  
Best Local Similarity 100.0%; Pred. No. 3e-108;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVFANTYVMIRGTPRNASTVSA 60  
DB 1 KRIMKVLGMIIVPYIMLFQWGRNRIIRFAGSLMALIVFANTYVMIRGTPRNASTVSA 60  
QY 61 TTSLVNSTNSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSKLRYPFTYGV 120  
DB 61 TTSLVNSTNSQVAAQEQNSSTSPAKHSTNSLOHQAHOAAATSSQSKLRYPFTYGV 120  
QY 121 GDLEIRVNSLOQVKSVDGIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180  
DB 121 GDLEIRVNSLOQVKSVDGIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLGNV 180  
QY 181 YQDSTAEIYANTNSGTIPDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240  
DB 181 YQDSTAEIYANTNSGTIPDLNPGVSWTTLVFDMPDFTYGVGHVGHYSLVASMGPFGS 240  
QY 241 DDTYALP 248  
DB 241 DDTYALP 248

#### RESULT 3

ID AEB45583  
AEB45583 standard; protein; 207 AA.

AC AEB45583;  
XX  
DT 22-SEP-2005 (first entry)

DE Alicyclobacillus sp. mature functional polypeptide (amino acids 42-248).  
XX Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.

```
XX OS Alicyclobacillus sp.; DSM 15716.
XX PN US2005147983-A.
XX PD 07-JUL-2005.
XX PF 23-FEB-2004; 2004US-00784592.
XX PR 06-JAN-2004; 2004DK-00000010.
XX PR 04-FEB-2004; 2004DK-00000165.
XX PA (NOVO ) NOVOZYMES AS.
XX PI Wilting R, Lassen SF, Ostergaard PR;
XX DR WPI; 2005-511773/52.
XX DR N-PSDB; AEB45582.
XX PT New functional polypeptides having function and amino acid sequence
XX PT similar to known specific bacterial enzymes useful in industrial,
XX PT research and household applications e.g. detergents and food.
XX PS Claim 36; Page; 83pp; English.
XX CC The present invention relates to functional polypeptides encoded by
XX CC polynucleotides comprised in the genome of Alicyclobacillus sp. DSM
XX CC 15716. The invention is useful in industrial, research and household
XX CC processes such as in composition to prepare feed and food, in detergent
XX CC formulations and for treating lignocellulosic fabric and pulp. The
XX CC present sequence is Alicyclobacillus sp. DSM 15716 mature functional
XX CC polypeptide. Note: this sequence is not shown in the specification but is
XX CC constructed based on the amino acid positions provided in claim 36 of the
XX CC specification.
XX SQ Sequence 207 AA;
Query Match 83.3%; Score 1074; DB 9; Length 207;
Best Local Similarity 100.0%; Pred. No. 7e-89;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 NTVMYKRGTPRNASTVSATSLVNSTSSQVAKOEONSTSPAHKSTNSLQHAHQAT 101
DB 1 NTVMYKRGTPRNASTVSATSLVNSTSSQVAKOEONSTSPAHKSTNSLQHAHQAT 60
QY 102 TSSSQSKLRYPFHTYGVKVDLEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDG 161
DB 61 TSSSQSKLRYPFHTYGVKVDLEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDG 120
QY 162 TMEVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTTPTDLNPGVSMTNLVFMDPDMT 221
DB 121 TMEVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTTPTDLNPGVSMTNLVFMDPDMT 180
QY 222 YGVHGOHYSLVASMGFSGDDETTALP 248
DB 181 YGVHGOHYSLVASMGFSGDDETTALP 207
RESULT 4
AAU37553
ID AAU37553 standard; protein; 331 AA.
XX AC AAU37553;
XX DT 14-FEB-2002 (first entry)
XX DB Staphylococcus aureus cellular proliferation protein #1723.
XX CC Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX CC antibacterial; drug design.
XX OS Staphylococcus aureus.
```

```
PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GD;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS5412.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 13146; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes,
XX CC themselves and the encoded proteins. The prokaryotes used are Baccharichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
Query Match 10.5%; Score 135; DB 4; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.0016;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
QY 115 HTYGV--TVGDLEIRVNSIQVKSQVYDGIGETANGAFVINITIRNDGTFMEVVDGIF 171
DB 81 HRTGVNKGDELVNVSVEVTKVSPSLAPNAGIFVADVITKNKGKALTTDSSMF 140
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTTPTD-----LNPGVSMTNLVFMDPDM 218
DB 141 KTKS-GDVTPEADNTGSMANGSDNGSIENSFTFLQRIKPNDSFAQKIVFDSB 192
RESULT 5
AAU34416
ID AAU34416 standard; protein; 331 AA.
XX AC AAU34416;
XX DT 14-FEB-2002 (first entry)
XX DB Staphylococcus aureus cellular proliferation protein #692.
XX CC Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX CC antibacterial; drug design.
```

```
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haeselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ,
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS52275.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 5912; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes,
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
XX
XX Query Match 10.5%; Score 135; DB 4; Length 331;
XX Best Local Similarity 31.0%; Pred. No. 0.0016;
XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
XX
XX QY 115 HTYG--KVGDLIRVNSLQVKSVGDIGETANGAFWVINITIRNDSTPMEVVDGIF 171
XX DB 81 HKIGTVNGDLIRVTVNSVETMKSVPSLAPTNAGIFVADVITIKNGKALTTIDSSMF 140
XX
XX QY 172 HLQNLNGNVYQPDSTAETIYAN-TNSGTPTD-----LNPQVSMNTNLVFDMPD 218
XX DB 141 KTKS-GDKTFEADNTGMSANQSDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192
XX
XX RESULT 6
XX ID AAU37279 standard; protein, 331 AA.
XX AC AAU37279;
XX XX 14-FEB-2002 (first entry)
XX XX Staphylococcus aureus cellular proliferation protein #1449.
XX DE
```

```
XX KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX KW antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US009180.
XX PR 21-MAR-2000; 2000US-0191078P.
XX PR 23-MAY-2000; 2000US-0206848P.
XX PR 26-MAY-2000; 2000US-0207727P.
XX PR 23-OCT-2000; 2000US-0242578P.
XX PR 27-NOV-2000; 2000US-0253625P.
XX PR 22-DEC-2000; 2000US-0257931P.
XX PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haeselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ,
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55138.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 12672; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 331 AA;
XX
XX Query Match 10.5%; Score 135; DB 4; Length 331;
XX Best Local Similarity 31.0%; Pred. No. 0.0016;
XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
XX
XX QY 115 HTYG--KVGDLIRVNSLQVKSVGDIGETANGAFWVINITIRNDSTPMEVVDGIF 171
XX DB 81 HKIGTVNGDLIRVTVNSVETMKSVPSLAPTNAGIFVADVITIKNGKALTTIDSSMF 140
XX
XX QY 172 HLQNLNGNVYQPDSTAETIYAN-TNSGTPTD-----LNPQVSMNTNLVFDMPD 218
XX DB 141 KTKS-GDKTFEADNTGMSANQSDNGSIENSFFLQRIINPDSTAQKIVFDVSE 192
XX
XX RESULT 7
XX ID ABU16122 standard; protein, 331 AA.
XX AC ABU16122;
XX XX ABU16122;
XX XX
```

DT 19-JUN-2003 (first entry)  
XX Protein encoded by prokaryotic essential gene #1649.  
DE  
XX  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
KM  
XX Staphylococcus aureus.  
OS  
XX  
XX WO20027183-A2.  
PN  
XX  
XX 03-OCT-2002.  
PD  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
PF  
XX  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342823P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
XX  
XX (ELIT-) ELITRA PHARM INC.  
PA  
XX Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW,  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,  
PI  
XX MPI: 2003-029926/02.  
DR N-PSDB; ACN19992.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX  
PS Claim 25; SEQ ID NO 44046; 1766bp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibody; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: the sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 331 AA;  
SQ

Query Match 10.5%; Score 135; DB 6; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.0016;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

115 HTYG--KVGDLEIRVNSLQGVKSVGYDGIETANGAFWVINTITRNDGSTMEVVDGIF 171

DB 81 HKIGETVANGDLEIRVNSLQGVKSVGYDGIETANGAFWVINTITRNDGSTMEVVDGIF 140  
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTIPTD-----LNDGVSMTNTLVFDMPD 218  
DB 141 KLSK-GDKTFEADNTGSMGANSQSDNGSIENSFPLQRIINPDSYAGKIVFDVSE 192  
RESULT 8  
ABM71142  
ID ABM71142 standard; protein; 350 AA.  
XX  
XX ABM71142;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Staphylococcus aureus protein #382.  
DE  
XX Staphylococcus aureus protein #382.  
XX  
XX Antimicrobial; vaccine; gene therapy; infection; sepsis; diagnosis;  
KM enzymatic assay; antibiotic target.  
XX  
XX Staphylococcus aureus.  
OS  
XX Staphylococcus aureus.  
XX  
XX WO200294868-A2.  
XX  
XX 28-NOV-2002.  
PD  
XX  
XX 27-MAR-2002; 2002WO-IB002637.  
PF  
XX  
XX 27-MAR-2001; 2001GB-00007661.  
PR  
XX  
XX (CHIR-) CHIRON SPA.  
PA  
XX  
XX Masignani V, Mora M, Scarselli M,  
PI  
XX MPI: 2003-120786/11.  
DR N-PSDB; ACP72702.  
XX  
XX New Staphylococcus aureus protein, useful as a vaccine for treating or  
PT preventing Staphylococcal infection, specifically an infection caused by  
PT *S. aureus*, e.g. sepsis.  
XX  
XX  
PS Claim 1; SEQ ID NO 764; 499p; English.  
XX  
XX The invention relates to novel genes and encoded proteins from  
CC Staphylococcus aureus. A composition comprising the *S. aureus* protein, a  
CC nucleic acid encoding the protein, or an antibody to the protein, is  
CC useful as a pharmaceutical, particularly as a vaccine for treating or  
CC preventing infection due to Staphylococcus bacteria, specifically an  
CC infection caused by *S. aureus*. The composition is particularly useful for  
CC treating or preventing sepsis in a patient. The composition can also be  
CC used for diagnostics. The protein is also used in an assay for enzymatic  
CC studies and as a target for antibiotics. This sequence represents one of  
CC the novel *S. aureus* proteins of the invention  
XX  
XX Sequence 350 AA;  
SQ

Query Match 10.5%; Score 135; DB 6; Length 350;  
Best Local Similarity 31.0%; Pred. No. 0.0017;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

115 HTYG--KVGDLEIRVNSLQGVKSVGYDGIETANGAFWVINTITRNDGSTMEVVDGIF 171  
DB 100 HKIGETVANGDLEIRVNSLQGVKSVGYDGIETANGAFWVINTITRNDGSTMEVVDGIF 159  
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTIPTD-----LNDGVSMTNTLVFDMPD 218  
DB 160 KLSK-GDKTFEADNTGSMGANSQSDNGSIENSFPLQRIINPDSYAGKIVFDVSE 211  
RESULT 9  
ABU15887  
ID ABU15887 standard; protein; 6713 AA.

```

XX AC ABU15887;
XX DT 19-JUN-2003 (first entry)
XX DB Protein encoded by Prokaryotic essential gene #1414.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Staphylococcus aureus.
XX PN WO20027183-A2.
XX PD 03-OCT-2002.
XX PE 21-MAR-2002; 2002MO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR N-PsDB; ACA19757.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 43811; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6713 AA;

```

Query Match 8.8%; Score 113.5; DB 6; Length 6713;  
 Best Local Similarity 24.4%; Pred. No. 8.3;

```

Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
QY 42 NTVMIRKNTPRNAST--VSATTSLVNSTNSQVAKQGNSTSPRAHSTNSLQAHQQA 99
DB 899 NONTYLDADSRKRNAYTQAVABGLNKQTCGNSTKADVNALNAVTKAKALNGASNLR 958
QY 100 ATSSSQSKLRYIPFHTYGVGDLLEIRVNSLQGVSVGYDGI---GETANQAFWYINITI 156
DB 959 NAKTSATNTINGLNLRLQLOQDNLKHQVEQAQNV--VGVNCKDKGNTLNTRMGALRTSI 1016
QY 157 RNDSTPMNEVVDGIFHLQNLNGVNYQPDSTAEIVANTNSGTIPDLNPGVSMNTLVVPM 216
DB 1017 QNDWTY--KTSQNTYLDASDSKNNY---NTA---VNNANGVYINATNMP--NMDANAINDM 1066

RESULT 10
AAU34143
ID AUJ4143 standard; protein; 2086 AA.
AC AAU34143;
DT 14-FEB-2002 (first entry)
DE Staphylococcus aureus cellular proliferation protein #419.
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;
KW antibacterial; drug design.
OS Staphylococcus aureus.
PN WO200170955-A2.
PD 27-SEP-2001.
PE 21-MAR-2001; 2001WO-US009180.
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zykkind JW, Wall D, Trawick JD, Carr GJ,
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PsDB; AAS52002.
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids.
XX PS Example 3; SEQ ID NO 5639; 511pp; English.
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the genes,
XX CC their use in the discovery of novel antibiotics, the essential genes
XX CC themselves and the encoded proteins. The prokaryotes used are Escherichia
XX CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
XX CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
XX CC useful for the identification of potential new targets for antibiotic
XX CC development. The antisense nucleic acids can also be used to identify
XX CC proteins used in proliferation, to express these proteins, and to obtain
XX CC antibodies capable of binding to the expressed proteins. The proteins can
XX CC be used to screen compounds in rational drug discovery programmes. The
XX CC antisense nucleic acid sequence is also useful to screen for homologous
XX CC nucleic acids which are required for cell proliferation in a wide variety
XX CC of organisms. The present sequence represents an essential prokaryotic
XX CC cellular proliferation protein. Note: The sequence data for this patent
XX CC did not form part of the printed specification, but was obtained in

```

```
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2086 AA;
Query Match 8.3%; Score 107.5; DB 4; Length 2086;
Best Local Similarity 24.2%; Pred. No. 5.9;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
Qy 42 NTVMIRGNTPRNAST--VSATTSLVNSTNSQVAKQKQNSSTSPAHKSTNSLQHQHQA 99
Db 1323 NQNTYLDADSKRNAYTQAVTAABGILNKQGTGNTSKADVNALMTVTAKALNGAENLR 1382
Qy 100 ATSSSQSKLRYIPHTYGVKVDLEIRVNSLQVKSQV-GYDGIETANGAFWVINITRN 158
Db 1383 NTKTSATNTINGLPNLTLQKDNLKHQVEQAVGVNGVDKQNTLTATAGALRTSIQ 1442
Qy 159 DGSIPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTIPTDLNPGVSMTTNLVFD 216
Db 1443 DNTT--KTSQNYLDASDSNKKNY--NTA---VNNANGVIVNTNPN--NMDANALNGM 1490
RESULT 11
ID AUA37017 standard; protein; 5795 AA.
XX
AC AUA37017;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #1187.
XX
KM Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (BLIT-) ELITRA PHARM INC.
XX
PI Haselebeck R, Ohlsen KU, Zyekind JW, Wall D, Trawick JD, Carr GJ,
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
XX
DR N-PSDB; AAS54876.
XX
PS New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 12610; 511P; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes,
CC themselves and the encoded proteins. The prokaryotes used are Bacteri
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
```

```
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: the sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5795 AA;
Query Match 8.3%; Score 107.5; DB 4; Length 5795;
Best Local Similarity 24.2%; Pred. No. 24;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
Qy 42 NTVMIRGNTPRNAST--VSATTSLVNSTNSQVAKQKQNSSTSPAHKSTNSLQHQHQA 99
Db 3600 NQNTYLDADSKRNAYTQAVTAABGILNKQGTGNTSKADVNALMTVTAKALNGAENLR 3659
Qy 100 ATSSSQSKLRYIPHTYGVKVDLEIRVNSLQVKSQV-GYDGIETANGAFWVINITRN 158
Db 3660 NTKTSATNTINGLPNLTLQKDNLKHQVEQAVGVNGVDKQNTLTATAGALRTSIQ 3719
Qy 159 DGSIPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTIPTDLNPGVSMTTNLVFD 216
Db 3720 DNTT--KTSQNYLDASDSNKKNY--NTA---VNNANGVIVNTNPN--NMDANALNGM 3767
RESULT 12
ID ABO70979 standard; protein; 417 AA.
XX
AC ABO70979;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #3154.
XX
KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX
DR N-PSDB; ABD04550.
XX
PS Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 19725; 4559P; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
```

CC production of *P. aeruginosa*-derived peptides or polypeptides, as target  
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused  
CC infection, and in detection of *P. aeruginosa* sequences or other sequences  
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-  
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format from USPTO at  
CC [www.uspto.gov/sequence.html](http://www.uspto.gov/sequence.html)  
XX Sequence 417 AA;  
SQ  
Query Match 8.3%; Score 106.5; DB 7; Length 417;  
Best Local Similarity 25.6%; Pred. No. 0.82;  
Matches 58; Conservative 33; Mismatches 79; Indels 57; Gaps 12;  
QY 32 GSWALIVFANTVVMIRGNTPRNASTVSATSLVNST-----NSSQVAKQEQNSSTSPA 67  
DB 21 GIKKALIVNTNTIASL---NTQRLNLSASASLNTSLQRLSTGRSRLNSKADPAQLQIANRL 77  
QY 68 ---TNSQVAKQEQNSSTSPA-----HKSTNSLQHAQH---QAATTSQSQSLRYIPF 114  
DB 78 TSGVNGALNVATKNNANDGISLQGTAEALQOSTNLTQMRDLSLQANGSNDSEKRL-- 135  
QY 115 HTGKRGVDELIRVNSLQOVKSVGYDGIGETANGAFWVINITIRNDGSTPEVVD-GIFHL 173  
DB 136 --NGEYKQLQ---KELDRISNTTTFGGRKLDGSEFGVASFGV--GSAANEIISVGDIDEM 187  
QY 174 --OVLNANVOPDSTABIYANTNSGTIPDINPGVS--MTNVLVFDM 216  
DB 188 SAELNLTGYFADGGAGVTAATNAGTV--DLAIGITGSAVNVAVDM 232  
RESULT 13  
AAU37120  
ID AAU37120 standard; protein: 2344 AA.  
XX  
AC AAU37120;  
XX  
DT 14-FEB-2002 (first entry)  
XX  
DE Staphylococcus aureus cellular proliferation protein #1280.  
XX  
KM Antisense; prokaryotic cellular proliferation protein; antibiotic;  
XX  
KW antibacterial; drug design.  
XX  
OS Staphylococcus aureus.  
XX  
FN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.  
XX  
PR 21-MAR-2000; 2000US-0191078P.  
XX  
PR 23-MAY-2000; 2000US-0206848P.  
XX  
PR 26-MAY-2000; 2000US-0207727P.  
XX  
PR 23-OCT-2000; 2000US-0242578P.  
XX  
PR 27-NOV-2000; 2000US-0253625P.  
XX  
PR 22-DEC-2000; 2000US-0257931P.  
XX  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KU, Zyekind JW, Wall D, Trawick JD, Carr GJ;  
XX Yamamoto RT, Xu HH;  
XX WPI; 2001-611495/70.  
XX DR N-PSDB; AAG54979.  
XX  
XX New polynucleotides for the identification and development of  
XX PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
XX Example 3; SEQ ID NO 12713; 511pp; English.

XX  
CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes,  
CC themselves and the encoded proteins. The prokaryotes used are *Bacteriella*  
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC [ftp.wipo.int/publ/published\\_pct\\_sequences](http://wipo.int/publ/published_pct_sequences)  
XX  
SQ Sequence 2344 AA;  
Query Match 8.2%; Score 106; DB 4; Length 2344;  
Best Local Similarity 25.9%; Pred. No. 9.5;  
Matches 59; Conservative 18; Mismatches 79; Indels 72; Gaps 9;  
QY 30 FAGSLMALIVFANTVVMIRGNTPRNASTVSATSLVNST-----NSSQVAKQEQNSSTSPA 85  
DB 89 FANSDAPLTSBELNTQSTTVGN--ONSTTIDASTADSTVTKNSSSV--QTSNSDVTSS 144  
QY 86 HKS-----TNSLQHAQHQAATTSQSQSLRYIPFHTYGVDELIRVNSLQOVKSVGYD 139  
DB 145 EKSENVSTSTNSNSNQEKLTSTSESTSK-----NTTSSDPTKSV--- 185  
QY 140 GIGETANGAFWVINITIRNDGSTPEVVDGIFHLQNLNANVOPDSTABIYANTNSGTIP 199  
DB 186 -----TSTSTEQPI-----NTSTNQSTAS--NNTSQSTTP 214  
QY 200 TDLNPGVSMNTNLVFDMPDFTYGVGHVGLSVASMGFPGSDERTYVL 247  
DB 215 TSANLMTKSTSTSTSTAPVKLTFPSRL-----ANSTFASATITTL 254  
RESULT 14  
ADJ34822  
ID ADJ34822 standard; protein: 922 AA.  
XX  
AC ADJ34822;  
XX  
DT 22-APR-2004 (first entry)  
XX  
DE Xylanase from an environmental sample seq id 38.  
XX  
KM antibacterial; fungicide; thermostable xylanase activity;  
XX  
KW dough conditioning; beverage production; nutritional supplement;  
XX  
KW animal feed; lignin reduction; wood product; xylan; bacterial infection;  
XX  
KW fungal infection; coccidiosis.  
XX  
OS Unidentified.  
XX  
FN WO2003106654-A2.  
XX  
PD 24-DEC-2003.  
XX  
PF 16-JUN-2003; 2003WO-US019153.  
XX  
PR 14-JUN-2002; 2002US-0389299P.  
XX  
PA (DIVE-) DIVERSA CORP.  
XX  
PI Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D;  
XX Bsteighlalian A;  
XX





**This Page Blank (uspto)**

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	173.5	13.5	2.5	2	H84032	hypothetical prote	
2	135	10.5	3.1	2	H89365	conserved hypothet	
3	120.5	9.3	5.2	2	S64030	probable membrane	
4	116.5	9.0	2.5	2	S77663	multicopy phage re	
5	113.5	8.8	67.3	2	B89921	hypothetical prote	
6	111	8.6	38.9	2	S53975	probable membrane	
7	102	7.9	3.5	2	B84443	hypothetical prote	
8	102	7.9	22.1	2	P90073	hypothetical prote	
9	101.5	7.8	3.4	2	A37853	flagellin, 40k - P	
10	100.5	7.8	2.2	2	B97130	uncharacterized se	
11	98	7.6	98.2	2	T13653	hypothetical prote	
12	97	7.5	8.6	2	AD2572	conserved hypothet	
13	97	7.5	8.8	2	B97454	hypothetical prote	
14	96.5	7.5	4.2	2	T40634	hypothetical prote	
15	96	7.4	4.3	2	S19377	probable membrane	
16	94	7.3	13.75	2	JT0345	dextranucrase (Eco	
17	93.5	7.3	4.8	1	FLCC	flagellin - Bacter	
18	93.5	7.3	7.4	2	S26638	SPR-1 protein - hu	
19	93.5	7.3	9.0	2	B87611	TonB-dependent rec	
20	92.5	7.2	12.8	2	A57384	multimerin, endothe	
21	91.5	7.1	2.1	2	G86678	hypothetical prote	
22	91.5	7.1	2.8	2	I64138	adhesin homolog HI	
23	91.5	7.1	14.9	2	T30552	glucosyltransferasa	
24	91.5	7.1	17.3	2	S30855	hypothetical prote	
25	90.5	7.0	2.3	2	H86925	probable lipoprote	
26	90.5	7.0	2.1	2	A99421	hypothetical prote	
27	90.5	7.0	13.65	2	T30822	limp protein - Myo	
28	90.5	7.0	18.0	2	AB1847	serine/threonine k	
29	90	7.0	3.44	2	T40167	hypothetical prote	

30	90	687	1	B65645	beta-galactosidase
31	90	865	2	T34584	probable secreted
32	90	1311	2	C84528	hypothetical prote
33	89.5	820	2	T17519	cell surface antig
34	89.5	1428	2	AC2224	hypothetical prote
35	89.5	1449	2	T30857	glucosyltransferasa
36	89.5	1475	2	B31135	gtf protein precu
37	89.5	1651	2	JC1340	outer membrane pr
38	89	385	2	T38113	hypothetical serin
39	89	453	2	T48240	hypothetical prote
40	88.5	365	2	D70043	hypothetical prote
41	88.5	511	2	C86798	prophage p13 prote
42	88.5	633	2	C32053	parasporel crystal
43	88.5	1770	2	S56221	hypothetical prote
44	88	116	2	B48338	hypothetical prote
45	88	1051	2	T18351	lmp1 protein - Myc

## ALIGNMENTS

RESULT 1  
 H84032  
 H84032  
 hypothetical protein BH3064 [imported] - Bacillus halodurans (strain C-125)  
 C|Species: Bacillus halodurans  
 C|Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
 C|Accession: H84032  
 R|Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiro  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A|Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A|Reference number: A83650; MUID:20512582; PMID:11058132  
 A|Accession: H84032  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-215 <STO>  
 A|Cross-References: UNIPROT:Q9R8E2; UNIPARC:UP100000C40B3; GB:AF001517; GB:BA000004; NIT  
 A|Experimental source: strain C-125  
 C|Genetics:  
 C|Gene: BH3064

Query Match	13.5%	Score 173.5;	DB 2;	Length 215;
Best Local Similarity	30.2%;	Pred. No. 3.2e-06;		
Matches 55;	Conservative 32;	Mismatches 80;	Indels 15;	Gaps 8

QY	69	NSSOVAAQ-EQNSSTSPHAKSTNSLQHQHQAAQ--TSSOSKALYIPHTYG---KYGD	122
Db	35	DSSTAOBPPEPBAEBSADQSENSSEBPBAEGTEEDTEBSAEEDDPIAGGALKVGD	94
QY	123	LEIRVNSLQQVSKSYGDIGETANGAFVAVITITINDGSTPEBREVDAIPEHLQNLNGV-Y	181
Db	95	VVFPAANGSTAGSVG-DVLTAEAKGTFILVDTVITINEGSDSLTYDSSPFKAK--VGDVEX	151
QY	182	QPDSTABIYANTNSGTFPTDINPGVSMITNLVFDMP-DFMTYGHVQGHYSILVASMGPFGS	240
Db	152	DSDSASAGIYANEGADFLTKLNPGLPELPGKVAVPDPVQLDSDDI----LNAVQGFPGT	207
QY	241	DE	242
Db	208	QQ	209

RESULT 2  
H89965  
conserved hypothetical protein SA1618 [imported] - *Staphylococcus aureus* (strain N315)  
C:Species: *Staphylococcus aureus*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: H89965  
R:Kurodo, M.; Ohea, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Matsuuchi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
Reference number: A89758; PMID:21311952; PMID:11418146



C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Oct-2004  
C/Accession: S53975  
R/Conor, R.; Churcher, C.M.  
submitted to the EMBL Data Library, April 1995  
A/Reference number: S53969  
A/Accession: S53975  
A/Molecule type: DNA  
A/Residues: 1-389 <CON->  
A/Cross-references: UNIPROT:Q04951; UNIPARC:UPI00001356B9; EMBL:Z49212; NID:g798940; PII  
C/Genetics:  
A/Gene: SGD:SCW10; MIPS:YMKR305C  
A/Cross-references: SGD:S0004921  
A/Map position: 13R  
C/Superfamily: family 17 glucosidase  
C/Keywords: transmembrane protein  
F/6-22/Domain: transmembrane #status predicted <TMM>

Query Match 8.6%; Score 111; DB 2; Length 389;  
Best Local Similarity 23.1%; Pred. No. 0.44;  
Matches 55; Conservative 40; Mismatches 95; Indels 48; Gaps 8;

QY 39 VPANTVYMRGNT-----PRNASTVSATTS-----LVNSTSSQVAKQFONS- 80  
DB 36 VHAQVTVVVSNGSGETIVNENAVVATTSRAVASQATTSTLEPTTSANVTSSQQQTST 95  
QY 81 -----STSPAHKSTNSLQHQHQQAATSSSQSK-LRIYIPHTYGVKVDLEI 125  
DB 96 LOSSEAASTVSGSTSSPSSSTSSASSASSASSASGALGITTSFPNDGSCCKSTQAQ 155  
QY 126 RVNSLQOVVSGVYDGIGETANGAFWYINITIRNDGSTPMWVDGIFHLQNLNGVYQPS 185  
DB 156 VASDLQGL--TGFDNRLXGVDCQVENVLQAKTSQKFL--GIYVVKIQDAVDITTS 211  
QY 186 TARIYANTNSGT--IPTDLNPGVSMTNLVPMDPDMTYGVGHYSLVAMGFFGS 240  
DB 212 AVESYGSWMDITTVSVGNELVNGSGSATTTQVGB-----YVSTAKSLTSGAYTGS 261

RESULT 7  
B84443  
Hypothetical protein At2G03000 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: B84443  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;  
euss, D.; Niernm, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: B84443  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-535 <STO->  
A/Cross-references: UNIPROT:O80614; UNIPARC:UPI000004AC2A; GB:AE002093; NID:g3461824; PII  
C/Genetics:  
A/Gene: At2G03000  
A/Map position: 2

Query Match 7.9%; Score 102; DB 2; Length 535;  
Best Local Similarity 24.3%; Pred. No. 3.3;  
Matches 53; Conservative 31; Mismatches 88; Indels 46; Gaps 8;

QY 18 FLOGMNRILRFAGSLMALIVFANTVYMRGNTPRNASTVSATTSLVNSTSSQVAKQF 77  
DB 277 YIQFGR-----LWE-----TNTT-----SSTRSRTTPTSTNVPSNSSSRVLQTS 317  
QY 78 QNS-STSPAHKSTN-SLQHQHQQAATSSSQSKLRIYIPHTYGVKVDLEIRVNSLQVKS 135  
DB 318 MSTRGTFPMSSSTRNSVQASMLAPGVSSSMSTVEIMPTSSARMTIMSEVANSFQTSY 377  
QY 136 VGVYDGIET-----ANG-----AFWVINTIRNDGSTPMWVDGIF 171

DB 378 IQFGLMTEFKVTYTLKSRVAVNENGEVCIQLEMRKEFTVYSTSNMRTSLMTLIRPID 437  
QY 172 HQNLNGVYQPSDSTARIYANTNSGTI--PTDLNPGVSM 208  
DB 438 EQISIRATVSSSTRTRARLLTNESPATIRAVAMLPVAM 475

RESULT 8  
F90073  
Hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)  
C/Species: Staphylococcus aureus  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: F90073  
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Chii, L.; Ogu-  
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratastu, K.  
Lancet 357, 1225-1240, 2001  
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A/Reference number: A69758; MUID:21311952; PMID:11418146  
A/Accession: F90073  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-2271 <KUR>  
A/Cross-references: UNIPROT:Q990Y4; UNIPARC:UPI00000CABB3; GB:BA000018; PII:g13702612; F  
A/Experimental source: strain N315  
C/Genetics:  
A/Gene: SA2447

Query Match 7.9%; Score 102; DB 2; Length 2271;  
Best Local Similarity 25.0%; Pred. No. 22;  
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVFANTVYMRGNTPRNASTVSATTSLVNST---NSQVAKQFONSTSPA 85  
DB 89 FAASDAPLTSBLNTQSEKTVGN--QNSTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 144  
QY 86 HKS-----TNSLQHQHQQAATSSSQSKLRIYIPHTYGVKVDLEIRVNSLQOVVSGVD 139  
DB 145 EKSEKVTSTVNSTSQOERLTSTSESTSSK-----NTSSSDTNSVA-- 186

QY 140 GIGETANGAFWYINITIRNDGSTPMWVDGIFHLQNLNGVYQPSDSTARIYANTNSGTIP 199  
DB 187 -----STSTREPT-----NSTNSTAS--NTSQSTIP 214

QY 200 TDLNPGVSMTNLVPMDPDMTYGVGHYSLVAMGFFGSDETYAL 247  
DB 215 SSVNLNKTSTSTSTPAVKLRTPSRL-----AMSTFASAAATTAV 254

RESULT 9  
A37853  
Flagellin, 40K - Pseudomonas aeruginosa  
C/Species: Pseudomonas aeruginosa  
C/Date: 28-Jun-1991 #sequence\_revision 28-Jun-1991 #text\_change 26-Aug-1999  
C/Accession: A37853  
R/Rotten, P.A.; Lory, S.  
J. Bacteriol. 172, 7188-7199, 1990  
A/Title: Characterization of the type A flagellin gene from Pseudomonas aeruginosa PAK.  
A/Reference number: A37853; MUID:91072275; PMID:2123866  
A/Accession: A37853  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-394 <TOT>  
A/Cross-references: UNIPARC:UPI000016FCA1; GB:M57501; NID:g151225; PII:NAAA63458.1; PII:  
C/Superfamily: flagellin

Query Match 7.9%; Score 101.5; DB 2; Length 394;  
Best Local Similarity 24.7%; Pred. No. 2.4;  
Matches 55; Conservative 33; Mismatches 78; Indels 57; Gaps 11;

QY 36 ALIVFANTVYMRGNTPRNASTVSA-----TTSLVNS-----T 68  
DB 2 ALTVNTIASL---NTQNLNNSASLNTSLQRLSTGSRINSAKDDAAGLOIANRLTGV 58

QY NSSQVAKOBNSSSTPA-----HKSTNSLOHAQH--QAATSSSQSLKRIYIPHTY 118  
DB 59 NGLNATNANDGISLQATBGAALQOSTNIIQRMEDLSLQANGSNSSSERALN----- 113  
QY 119 KVGDELRVNSLQOVKSVGDIGETANGAFWVINTIRNDSTMEVVD-GIFHL--QN 175  
DB 114 --GEAKQLOKELDRISNTTFGGRKLDSFGVASFQV---GSANREILSVGIDMSAES 168  
QY 176 LINGNVQPDSTAEIYANTNSGTFPTDLPNGV--MTNVLVPRM 216  
DB 169 LNTGYFKADGGGAVTAATASGTV--DIALGITGSAAVNVKVM 209

RESULT 10  
E97130  
uncharacterized secreted protein, homolog YXK Bacillus subtilis [imported] - Clostridium  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004  
C/Accession: E97130  
R.Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium CLO  
A/Reference number: A96900; MUID:21359325; PMID:21359325  
A/Accession: E97130  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-202 <KUR>  
A/Cross-references: UNIPROT:Q97HY8; UNIPARC:UP100000CA32C; GB:AE01437; PIDN:AAK79832.1;  
A/Experimental source: Clostridium acetobutylicum ATCC824  
C/Genetics:  
A/Gene: CAC1868

Query Match 7.8%; Score 100.5; DB 2; Length 202;  
Best Local Similarity 23.5%; Pred. No. 1.2;  
Matches 40; Conservative 38; Mismatches 71; Indels 21; Gaps 6;  
QY 85 AHKSTNSLOHAQHQAATSSSQSLKRIYIPHTYGVGLERVNSLQOVKSV-GYDGI 142  
DB 33 AKETTSATKVVQNNKATSTDKQDKIKYVGBGSKGAMSIKVDLTQETNTIAGDGEN 92  
QY 143 ETANGAFWVINTIRNDSTPMENVGIFHLQNLGNV-YQPDSTA-----EIVANT 193  
DB 93 KTTQOKFFVVVHLQMTNKENNVAGYBPDNLDGDIKTKAGYKMDMEGETANGAKITVAKD 152  
QY 194 NS-GITPTDLPNGVSMNTMLVFDMPDPMYTHVGVGHYSLVASMGFFGSDE 242  
DB 153 DSFFGVYDKVNPMLSKQYIVFEVP-----TNFINANAVLTHGGDD 193

RESULT 11  
T13653  
hypothetical protein 95B7.2 - fruit fly (Drosophila melanogaster)  
C/Species: Drosophila melanogaster  
C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
C/Accession: T13653  
R.Perez, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.  
Submitted to the EMBL Data Library, April 1999  
A/Description: Sequencing the distal X chromosome of Drosophila melanogaster.  
A/Reference number: Z17694  
A/Accession: T13653  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-962 <FER>  
A/Cross-references: UNIPARC:UP10000128461; EMBL:AL021728; NID:e1355643; PID:e1301389; PI  
C/Genetics:  
A/Cross-references: FlyBase:FBgn0000376  
A/Introns: 181/1; 249/1; 774/3; 899/1; 949/3  
A/Note: RG:95B7.2

Query Match 7.6%; Score 98; DB 2; Length 982;  
Best Local Similarity 20.3%; Pred. No. 15;

Matches 44; Conservative 44; Mismatches 89; Indels 40; Gaps 8;  
QY 17 LFIQGNRNR-LRFAGSLALI-----VFANTVYMGNTPRNASTVATSLVNSTSS 71  
DB 596 LTVFQDSKLELR--YWCEDVSTAMASILITITVTSSSVGTOTSSASNAOT 652  
QY 72 QVAKOBNSSSTPAHKSTNSLOHAQH-----QAATSSSQSLKRIYIPF----- 114  
DB 653 SASNCVYASASSNSTSTSLPYNPDCDSYERAVATVTSVSNLKLILLVANLSERV 712  
QY 115 -----HTYKVGDLERVNSLQOVKSVGDIGETANGAFWVINTIRNDSTMEV 167  
DB 713 KRCNCGHTCDKRCRDLMTKAQOLAEATYV--GVGVVGRFRTPMLFVR--PISNI 766  
QY 168 DGIFHLQNLGNVQPDSTAE--IYANTNSGTFPTD 202  
DB 767 DVPRLSLMTROKISROVLVORLLPPTSADRPYL 803

RESULT 12  
AD2672  
conserved hypothetical protein Atu0778 [imported] - Agrobacterium tumefaciens (strain C5  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C/Accession: AD2672  
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McChell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krepan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A/Reference number: AB2577; MUID:21608550; PMID:11743193  
A/Accession: AD2672  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-846 <KUR>  
A/Cross-references: UNIPROT:Q8UH45; UNIPARC:UP10000164523; GB:AE008688; PIDN:AAL41794.1;  
A/Experimental source: strain C58 (Dupont)  
C/Genetics:  
A/Gene: Atu0778  
A/Map position: circular chromosome

Query Match 7.5%; Score 97; DB 2; Length 846;  
Best Local Similarity 21.4%; Pred. No. 14;  
Matches 62; Conservative 30; Mismatches 100; Indels 98; Gaps 12;  
QY 48 RGNTPRNASTVATSLVNSTNSQVAKO-----EONSSSTPAHKSTNSLOHAQH 98  
DB 458 RGIDPTNRAVISTPSSAGSTSSPSASASPSGSSNTSTPATVTSGAASTPTAS 517  
QY 99 AATTSSSQSLKRIYIPHTYGVGLERVNSLQOVKSVGDIGETANGAFWVINTIRNDSTMEVVD-GIFHL--QN 175  
DB 518 SVSSPSGSAKKGHSILPARSGIIDDQY--AGTAAARNAGVGINIANIVLDGAGMPMR 575  
QY 129 -----SLOOVKSVGDIGETANGAFWVINTIRNDSTMEVVDGIFHLQNL 177  
DB 576 SLIEGRQSPFSLVSVGNND-GRDAAGAFGLADIGYGLDGGATARAFAFGLYDQRIN 634  
QY 178 -----GNYQPDSTAEIYANTNSGTFPTDLPNGV--MTNVLVPRM 216  
DB 635 TCGDFTHSGFYIAPRISLPLTDGLVATIGYVYAPGRMSIERGYLNGGTMDYSRGETDIN- 693  
QY 205 GVSMTNVLVFPMPDPMYTG-----HYQGHYSLVASMGFFGSDETTVALP 248  
DB 694 --VMAKLRFLWDLALTTIGENMLTPYAGVYA-KAEMGAYV--ETGGAFP 738

RESULT 13  
B97454  
hypothetical protein AGR\_C1421 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C/Species: Agrobacterium tumefaciens  
C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004

C:Accession: B97454  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Goldman, A.; Liu, F.; Woliam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B. Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*  
A:Reference number: A87359; PMID:11743194  
A:Accession: B97454  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-848 <KUR>  
A:Cross-references: UNIPROT:O8UHA5; UNIPARC:UPI00000D1910; GB:AE007869; PIDN:AAK6587.1, C:Genetics:  
A:Gene: AGR\_C\_1421  
A:Map position: circular chromosome

Query Match	7.5%;	Score 97;	DB 2;	Length 848;
Best Local Similarity	21.4%;	Pred. No. 15;		
Matches	62;	Conservative	30;	Mismatches 100;
				Indels 98;
				Gaps 12

```

QY      48 RNTPRNASTVATATSTLVNVTSSQVAKQ-----EONSSTPAHKSTSLQHAQHQ 98
Db      460 RQIDPTNRNRAVSTTTPBASSGTSSSPSSSSASABSSGSGSNTSTSPATTTSGAASTPTAS 51.9
QY      99 AATTSSSGCKLR-----IPHTYKGVQDLIRVN-----128
Db      520 SVSSPQSAKTKGSHILPARGIILDIQY--ACTIARPNAGVAGTANIVLDAGHGMPMR 577
QY      129 -----SLQQYKSVGYDGIQGTANGAFVNIIT---RNDGSTPHEVVDGI.FH.LQNIN 177
Db      578 SLIBPGRQSFIVSDVGHND-GRDAAGAAGIADIGYGFGLDGGATARIAPAGLYDQKDIN 636
QY      178 -----GNVYQPDSTARIYANTNGTI-----PTDLNP 204
Db      637 TCGDFPIHSGFYIAPBISLPLTQGLVANTIGGYVAPGMSISERGVLNGTMDYSRGETDLN- 695
QY      205 GYSMTNINLVDPMDPENTYQ-----HYGQHSYLSVAMSGFFGSDETTYALP 248
Db      636 --VMAAKRFMDLDAITIGEMNLTTPAGATYA--KAAMGANA--ETGAPF 740

```

## RESULT 14

hypothetical protein SPBC685.03 - fission yeast (*Schizosaccharomyces pombe*)  
 C.Species: Schizosaccharomyces pombe  
 C.Date: 03-Dec-1999 #sequence 03-Dec-1999 #text\_change 09-Jul-2004  
 C.Accession: T40634  
 R.Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, March 1999  
 A.Reference number: Z21928  
 A.Accession: T40634  
 A.Status: preliminary; translated from GB/EMBL/DBJ  
 A.Molecule type: DNA  
 A.Residues: 1-452 <SEE>  
 A.Cross-references: UNIPROT:Q9VYL5; UNIPARC:UPI000006ACFA; EMBL:AL049474; PIDN:CAB39360  
 A.Experimental source: strain 972h-, cosmid c685  
 C.Genetics:  
 A.Gene: SPDB:SPBC685.03  
 A.Map position: 2

Query Match	7.5%	Score 96.5;	DB 2;	Length 452;
Best Local Similarity	23.2%;	Pred. No. 6.9;		
Matches	54;	Conservative	36;	Mismatches 92;
				Indels 51;
				Gaps 8;

[illegible]

Db 225 SVYNNGSA-VSNVTYVNTKMTSSNKKCCYDDVYVANI FGLDFTTAAVLSBVSILBSFALCNA 283

Gy 196 GTTPTDLNPGVSMITNLVFDMPDFTYTG-----HVGQHSLSVASMGPFGSSD 241

Db 284 TT-----SSSLFRQGLASVGVGVSPHPSSESSGFANLLGTNNYFMTD 325

## RESULT 15

C:Probable membrane protein YCL048w - Yeast (*Saccharomyces cerevisiae*)  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 05-Oct-2004  
 C:Accession: S19377  
 R:Grenson, M., Jauniaux, J.C.; Ureastarazu, I.A.  
 submitted to the Protein Sequence Database, March 1992  
 A:Reference number: S19376  
 A:Accession: S19377  
 A:Molecule type: DNA  
 A:Residues: 1-463 <GRE>  
 A:Cross-references: UNIPROT:P25380; UNIPARC:UP100001345FC; EMBL:X59720; NID:g1907116; P  
 C:Genetics:  
 A:Gene: MIPS:YCL048w  
 A:Cross-references: SGD:S0000553  
 A:Map position: 3L  
 C:Superfamily: sporulation-specific protein 2 precursor  
 C:Keywords: transmembrane protein  
 F7-79/Domain: transmembrane #status predicted <TM1>  
 F146-463/Domain: transmembrane #status predicted <TM2>

Query Match	7.4%;	Score 96;	DB 2;	Length 463;
Best Local Similarity	22.2%;	Pred. NO. 7.8;		
Matches 51;	Conservative 33;	Mismatches 80;	Indels 66;	Gaps 9

```

OY 16 MELFIQGRNRI-----LRFGSI-VALLIYFANTVMIRGNTFRNA 55
Db 130 LIRIÖGNKJAKRGLEFQLOGLTSLVSVEIPTLKFCQSLMKVVPILIVYSDSONIEI IK 1898
OY 56 STVSATTSLVNSTNSOVAKOEONSTSPAKSTNSLOHAQOATTSOSSOKURYPIH 115
Db 190 DIYDPTSLANIEHPKV--QSIDT---FMINNRPLETHSNVKTIIRQ-----FS 236
OY 116 TYGRVDDLEIRVNSLOQVKSVDIGETANGAFWVINTIRNDG-----STPME 1655
Db 237 VHANAKELEIEMPHLEKVS-----NITIRDSLVLPLQITVKXSLK 278
OY 166 VVDGIPIHLÖNGNGVYQPDSTAEIYANTNSGIP---TDLNPVSMITN 211
Db 279 FIENYVEIENLN-NÖKIGTGLGINNVNLKIVNLENITLIDÖGGLMADN 327

```

Search completed: April 7, 2006, 16:24:07  
Job time : 30.2527 secs

This Page Blank (uspto)



## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

This Page Blank (uspio)

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:09:09 ; Search time 162.426 Seconds

(without alignments)  
1077.233 Million cell updates/sec

Title: US-10-784-592-43

Perfect score: 1289

Sequence: 1 MRIMKVLGMLVPIYIMLFIQ.....YSLVSMGPFSGSDFTTALP 248

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 05.80:\*

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	173.5	13.5	215	2	09K8E2_BACHD
2	139	10.8	331	2	06G8D3_STPAS
3	139	10.8	331	2	08NVZ1_STPAW
4	135	10.5	331	2	06G8Q7_STPAR
5	135	10.5	331	2	05HEX9_STPAC
6	135	10.5	331	2	07A4Z0_STPAM
7	135	10.5	331	2	09PT70_STPAM
8	122	9.5	202	2	04MIR5_BACCB
9	120.5	9.3	542	1	SCW11_YEAST
10	119.5	9.3	170	2	08KB32_CHLRE
11	116.5	9.0	215	2	050442_MYCSM
12	113.5	8.8	6713	2	0931R6_STPAM
13	113.5	8.8	6713	2	093U54_STPAM
14	111.5	8.7	387	1	086265_PSEAB
15	111	8.6	389	1	SCW10_YEAST
16	108.5	8.4	229	2	09ZGA7_CLOCE
17	106.5	8.3	10746	2	06G8A3_STPAR
18	105.5	8.2	329	2	06SYM1_PASMU
19	105.5	8.2	387	2	086266_PSEAB
20	105.5	8.2	393	1	FLICA_PSEAB
21	105.5	8.2	394	2	0532R9_PSEAB
22	105.5	8.2	394	2	0532S1_PSEAB
23	105	8.1	2370	2	082RB3_STPAM
24	104.5	8.1	766	2	054PB6_DICDI
25	104	8.1	9804	2	08RWOC_STPAM
26	103.5	8.0	923	2	07UKM5_RHOBA
27	103	8.0	713	2	06FE57_ACTAD
28	102.5	8.0	1259	2	0869X4_DICDI
29	102	7.9	535	2	080614_ARATH
30	102	7.9	2261	2	05HCP3_STPAC
31	102	7.9	2271	2	07A362_STPAM

32	102	7.9	2271	2	0990Y4_STPAM	0990Y4 staphylococ
33	102	7.9	2275	2	06G620_STPAS	06G620 staphylococ
34	102	7.9	2275	2	08NVZ1_STPAW	08NVZ1 staphylococ
35	101.5	7.9	1272	2	04UDM8_TREAN	04UDM8 thelteria a
36	100.5	7.8	202	2	097HY8_CLOAB	097HY8 clostridium
37	100	7.8	1915	2	09RPL0_9CLOT	09RPL0 acetivibrio
38	100	7.8	2283	2	08V099_STPAU	08V099 staphylococ
39	99.5	7.7	696	2	054Y58_DICDI	054Y58 dictyostell
40	99	7.7	285	2	04WL74_ASDFU	04WL74 aspergillius
41	99	7.7	333	2	P72123_PSEAB	P72123 pseudomonas
42	99	7.7	340	2	P72123_PSEAB	P72123 pseudomonas
43	99	7.7	355	2	05Z159_NOCFA	05Z159 nocardia fa
44	99	7.7	1221	2	054IUB_DICDI	054IUB dictyostell
45	99	7.7	1564	2	069237_MOUSE	069237 mus musculu

## ALIGNMENTS

RESULT 1  
09K8E2\_BACHD  
ID 09K8E2\_BACHD PRELIMINARY; PRT; 215 AA.  
AC 09K8E2;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)  
DR BH3064 protein.  
GN OrderedLocustNames=BH3064;  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
OX NCBI\_Taxid=86665;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;  
RA Takami H., Nakasone K., Takaki Y., Mieno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331 (2000).  
DR EMBL; BA000004; BAB06783.1; -; Genomic\_DNA.  
DR PIR; H84032; H84032.  
KW Complete proteome.  
SQ SEQUENCE 215 AA; 23072 MW; 6A05DA5DE7DC6358 CRC64;

Query Match 13.5%; Score 173.5; DB 2; Length 215;  
Beet Local Similarity 30.2%; Pred. No. 9.5e-06;  
Matches 55; Conservative 32; Mismatches 80; Indels 15; Gaps 8;  
QY 69 NSSQVAKQ-EONSSTSPARKSTNSLOHQAQAA--TSSSOKLRYIPFTYG---KVGD 122  
DB 35 DSSSTAGQPEPBAEVSADQSENESEPEBAEPTEDTBEEBAEEDPDIAGGALTKVD 94  
QY 123 LEIRVNSIQQVKSVDGIGETANGAFVIVITIRNDGSTPEVVDGIFHLQNLGNV-Y 181  
DB 95 VVFETANGSTAGSVG-DVLTAAKGTFLIVDTIKNBSDSITVDSFFPKK--VGDEY 151  
QY 182 QPDSTAAETIYANTNSGITITDINPGVSMITNLVFDMP-DPMTYGVGQHTSYLVASNGFSGS 240  
DB 152 DSDSSAGLVANEGADFFLTKLINPGLLEPGKVFDPQDVLSDDI----LINVQNGFFGT 207  
QY 241 DE 242  
DB 208 QQ 209  
RESULT 2  
06G8D3\_STPAS  
ID 06G8D3\_STPAS PRELIMINARY; PRT; 331 AA.  
AC 06G8D3;  
DT 05-JUL-2004 (TREMblrel. 27, Created)  
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
DE Putative membrane protein.  
GN OrderedLocustNames=SA51720;  
OS Staphylococcus aureus (strain MSSA476).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=282459;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;  
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,  
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,  
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,  
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,  
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,  
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,  
RA Spratt B.G., Parkhill J.;  
RT "Complete genomes of two clinical Staphylococcus aureus strains:  
RT evidence for the rapid evolution of virulence and drug resistance.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
DR EMBL: BX571857; CAG43524.1; -; Genomic\_DNA.  
KW Complete proteome.  
SQ SEQUENCE 331 AA; 35855 MW; 753AF5FDAF50CF9 CRC64;  
Query Match 10.8%; Score 139; DB 2; Length 331;  
Best Local Similarity 31.9%; Pred. No. 0.0094;  
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;  
QY 115 HTYG---KVGDLEIRVNSLQGVKSVGYDGETANGAFWVINITIRNDGSTPMEVVDGIF 171  
DB 81 HKIGETVNGDLEVTAVNSVETKSVGPSIAPTNAGTIVADVITKNGKALITDSSMF 140  
QY 172 HLOINANGVYOPDSTARIYANTN-SGTIPTD-----LNPGVSMNTNLVFDMPD 218  
DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192  
RESULT 3  
ID OSNVZ1\_STAAM PRELIMINARY; PRT; 331 AA.  
AC OSNVZ1,  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DE 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
GN Hypothetical protein MM1738.  
OS OrderedLocustNames=MM1738;  
OC Staphylococcus aureus (strain MW2).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=196620;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;  
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,  
RA Yamamoto K., Hizematsu K.;  
RT "Genome and virulence determinants of high virulence community-  
RT acquired MRSA.";  
RL Lancet 359:1819-1827(2002).  
DR EMBL: BA000033; BAB95603.1; -; Genomic\_DNA.  
KW Complete proteome; Hypochemical protein.  
SQ SEQUENCE 331 AA; 35855 MW; 753AF5FDAF50CF9 CRC64;  
Query Match 10.8%; Score 139; DB 2; Length 331;  
Best Local Similarity 31.9%; Pred. No. 0.0094;  
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;  
QY 115 HTYG---KVGDLEIRVNSLQGVKSVGYDGETANGAFWVINITIRNDGSTPMEVVDGIF 171  
DB 81 HKIGETVNGDLEVTAVNSVETKSVGPSIAPTNAGTIVADVITKNGKALITDSSMF 140  
QY 172 HLOINANGVYOPDSTARIYANTN-SGTIPTD-----LNPGVSMNTNLVFDMPD 218

DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192  
RESULT 4  
ID OS6F07\_STAAR PRELIMINARY; PRT; 331 AA.  
AC OS6F07,  
DT 05-JUL-2004 (TReMBLrel. 27, Created)  
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)  
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)  
GN Putative membrane protein.  
OS OrderedLocustNames=SA1880;  
OC Staphylococcus aureus (strain MRSA252).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=282459;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;  
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,  
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,  
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,  
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,  
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,  
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,  
RA Spratt B.G., Parkhill J.;  
RT "Complete genomes of two clinical Staphylococcus aureus strains:  
RT evidence for the rapid evolution of virulence and drug resistance.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
DR EMBL: BX571856; CAG40870.1; -; Genomic\_DNA.  
KW Complete proteome.  
SQ SEQUENCE 331 AA; 35850 MW; 82E9C361D2D59066 CRC64;  
Query Match 10.5%; Score 135; DB 2; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.02;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;  
QY 115 HTYG---KVGDLEIRVNSLQGVKSVGYDGETANGAFWVINITIRNDGSTPMEVVDGIF 171  
DB 81 HKIGETVNGDLEVTAVNSVETKSVGPSIAPTNAGTIVADVITKNGKALITDSSMF 140  
QY 172 HLOINANGVYOPDSTARIYAN-TNSGTIPTD-----LNPGVSMNTNLVFDMPD 218  
DB 141 KTKS-GDKTFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192  
RESULT 5  
ID OSHEX9\_STAAC PRELIMINARY; PRT; 331 AA.  
AC OSHEX9,  
DT 10-MAY-2005 (TReMBLrel. 30, Created)  
DT 10-MAY-2005 (TReMBLrel. 30, Last sequence update)  
DE 10-MAY-2005 (TReMBLrel. 30, Last annotation update)  
GN Conserved domain protein, putative.  
OS OrderedLocustNames=SA001847;  
OC Staphylococcus aureus (strain COU).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=93062;  
RN [1]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX PubMed=15774886; DOI=10.1126/JB.187.7.2426-2438.2005;  
RA Gill S.R., Fouts D.E., Archer G.L., Mongodin E.F., DeBoy R.T.,  
RA Ravel U., Paulsen I.T., Kolonay J.F., Brinkac L.M., Beaman M.J.,  
RA Dodson R.J., Daugherty S.C., Madupu R., Angiuoli S.V., Durkin A.S.,  
RA Hatt D.H., Vamathevan J.J., Khouri H., Uetaback T.R., Lee C.,  
RA Dimetrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H.,  
RA Hance I.R., Nelson K.E., Fraser C.M.;  
RT "Insights on evolution of virulence and resistance from the complete  
RT genome analysis of an early methicillin-resistant Staphylococcus  
RT aureus strain and a biofilm-producing methicillin-resistant  
RT Staphylococcus epidermidis strain.";  
RL J. Bacteriol. 187:2426-2438(2005).

DR EMBL; CP000046; AAM36865.1; -; Genomic\_DNA.  
DR TIGR; SACL01847; -  
KW Complete proteome.  
SQ SEQUENCE 331 AA; 35864 MW; F4347361DD259472 CRC64;

Query Match 10.5%; Score 135; DB 2; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.02;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG---KVGDLIRVNSIQOVKSVGYDGIGETANGAFWVINTITRNDSTPMEVVDGIF 171  
DB 81 HKIGETVKNKGDELVTVNSVETMKSVGPSLAPTNAGIFVAVDTITKNGKEALTIDSSMF 140  
QY 172 HQINANGVYQDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218  
DB 141 KIKS-GDKTFEADNTGMSANQSDNGSIENSFPFLQRIINPDSIAOGKIVFDVSE 192

## RESULT 6

Q7A4Z0 STAAAN PRELIMINARY; PRT; 331 AA.

AC Q7A4Z0;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein SAI618.  
GN OrderedLocustNames=SAI618.  
OS Staphylococcus aureus (strain N315).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxId=158879;  
RN [1]

RP NUCLEOTIDE SEQUENCE [Large scale genomic DNA].

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiratsuku K.,  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus."  
RL Lancet 357:1225-1240(2001).  
DR EMBL; BA000018; BAB4286.1; -; Genomic\_DNA.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 331 AA; 35892 MW; F25DA94971A1AAF CRC64;

Query Match 10.5%; Score 135; DB 2; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.02;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG---KVGDLIRVNSIQOVKSVGYDGIGETANGAFWVINTITRNDSTPMEVVDGIF 171  
DB 81 HKIGETVKNKGDELVTVNSVETMKSVGPSLAPTNAGIFVAVDTITKNGKEALTIDSSMF 140  
QY 172 HQINANGVYQDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218  
DB 141 KIKS-GDKTFEADNTGMSANQSDNGSIENSFPFLQRIINPDSIAOGKIVFDVSE 192

## RESULT 7

Q99T70 STAAAN PRELIMINARY; PRT; 331 AA.

AC Q99T70;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein.  
GN OrderedLocustNames=SAV1800;  
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxId=158878;  
RN [1]

RP NUCLEOTIDE SEQUENCE [Large scale genomic DNA].

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiratsuku K.,  
RT "Whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus."  
RL Lancet 357:1225-1240(2001).  
DR EMBL; BA000017; BAB57962.1; -; Genomic\_DNA.  
DR PIR; H89965; H89965.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 331 AA; 35892 MW; F25DA94971A1AAF CRC64;

## Query Match

10.5%; Score 135; DB 2; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.02;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG---KVGDLIRVNSIQOVKSVGYDGIGETANGAFWVINTITRNDSTPMEVVDGIF 171  
DB 81 HKIGETVKNKGDELVTVNSVETMKSVGPSLAPTNAGIFVAVDTITKNGKEALTIDSSMF 140  
QY 172 HQINANGVYQDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218  
DB 141 KIKS-GDKTFEADNTGMSANQSDNGSIENSFPFLQRIINPDSIAOGKIVFDVSE 192

## RESULT 8

Q4MIR5 BACCE PRELIMINARY; PRT; 202 AA.

ID Q4MIR5;  
AC Q4MIR5;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein.  
GN ORFNames=BCE\_G9241\_PBClin29\_0020;  
OS Bacillus cereus G9241.  
OC Bacteria; Firmicutes; Bacillales; Bacillus;  
OC Bacillus cereus group.  
OX NCBI\_TaxId=269801;  
RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=G9241;  
RX PubMed=15155910; DOI=10.1073/pnas.0402414101;  
RA Hoffmaster A.R., Ravel J., Raeko D.A., Chapman G.D., Chute M.D.,  
RA Marston C.K., De B.K., Sacchi C.T., Fitzgerald C., Mayer L.W.,  
RA Maiden M.C.J., Priest F.G., Barker M., Jiang L., Cer R.Z.,  
RA Ralston J., Fraser C.M., Weyant R.S., Galloway D.R., Read T.D.,  
RA Popovic T., Pater S.N.,  
RT "Identification of anthrax toxin genes in a Bacillus cereus associated  
RT with an illness resembling inhalation anthrax."  
RL Proc. Natl. Acad. Sci. U.S.A. 101:8449-8454(2004).  
CC -!- CAUTION: The sequence shown here is derived from an  
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is  
CC preliminary data.  
DR EMBL; AA00100053; EAL12062.1; -; Genomic\_DNA.  
KW Hypothetical protein.  
SQ SEQUENCE 202 AA; 22430 MW; 7AA489138004FE0C CRC64;

Query Match 9.5%; Score 122; DB 2; Length 202;  
Best Local Similarity 23.5%; Pred. No. 0.11;  
Matches 44; Conservative 27; Mismatches 94; Indels 22; Gaps 4;

QY 35 WALIVPANTYVIMIKNTIRANSTYSATTSVNSSTSGVAQEQNSSTSPAKSTNSIOH 94  
DB 10 WFWIVIVIVVIGAYGN-----STDKERTKTASTEPKQEAQOE-TRKDEPKPEKASBP 63  
QY 95 AOHQAATSSQSKLRVIFPHYGVKVGDLIRVNSIQOVKSVGYDGIGETANGAFWVINTI 154  
DB 64 KEPPKEIKSKE-----GESSKVKIAGSVSESTDSVGQYLSEKAGQFKVVEI 111

```

QY 155 TIRNDGSTRMEVVDGIFHLQNLNGVYQPDSTAETIYANTNSG-----TIPTDLNPGVSWTT 210
DB 112 SITNNQKAIITVDANSFTLVVNNQDEFRKSTQAGAFVNGKSKDFPLKQNLNPLTQTGT 171
QY 211 NLVDPMP 217
DB 172 KIIFDVP 178

RESULT 9
SCW11_YEAST STANDARD; PRT; 542 AA.
AC PS3189;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Probable family 17 glucosidase SCW11 precursor (EC 3.2.1.-) (Soluble
cell wall protein 11).
GN Name=SCW11; OrderedLocNames=YGL028C;
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OK NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=97313265; PubMed=9169669;
RA Tetteijn H., Agoetoni-Carbone M.L., Albermann K., Albers M.,
RA Arroyo J., Backes U., Barreiro T., Bertani I., Bjournon A.J.,
RA Bruecker M., Brusch C.V., Cartignat G., Castagnoli L., Cerdan E.,
RA Clemente M.L., Coblenz A., Coglieva M., Colasac E., Defoor E.,
RA Del Bino S., Delius H., Delneri D., de Wergifosse P., Dujon B.,
RA Durand P., Eutian K.-D., Eraso P., Sacrihano V., Fabiani L.,
RA Fattmann B., Feroli F., Feuermann M., Frontali L., Garcia-Gonzalez M.,
RA Garcia-Saez M.I., Goffeau A., Guerreiro P., Hami J., Hansen M.,
RA Hebling U., Hernandez K., Heumann K., Hilger F., Hofmann B.,
RA Indge K.J., James C.M., Klima R., Koetter P., Kramer B., Kramer W.,
RA Lauguin G., Leuther H., Louis E.J., Maillyer E., Marconi A.,
RA Martegani E., Mazon M.J., Mazzoni C., McReynolds A.D.K.,
RA Melchiorreto P., Mewes H.-W., Minkova O., Mueller-Aer S.,
RA Nawrocki A., Nötter P., Neu R., Nombela C., Oliver S.G., Panzeri L.,
RA Paoluzzi S., Plevani P., Portetelle D., Portillo F., Potier S.,
RA Purnelle B., Rieger M., Riles L., Rinaldi T., Robben J.,
RA Rodrigues-Pouzada C., Rodriguez-Belmonte E., Rodriguez-Torres A.M.,
RA Rose M., Ruzzi M., Sallola M., Sanchez-Perez M., Schaefer B.,
RA Souciet J.-L., Steenma H.Y., Talla E., Thierry A., Vandembol M.,
RA van der Aart Q.J.M., Van Dyck L., Vanoni M., Verhasselt P., Voet M.,
RA Volckaert G., Wambut R., Watson M.D., Weber N., Wedler E., Wedler H.,
RA Wipfli P., Wolf K., Wright L.F., Zaccaria P., Zimmermann M.,
RA Zollner A., Kleine K.;
RL "The nucleotide sequence of Saccharomyces cerevisiae chromosome VII.",
RL Nature 387:81-84(1997).
RN [2]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RC STRAIN=5288c / SEY6210;
RX MEDLINE=98422453; PubMed=9748433;
RA Cappellaro C., Moya V., Tanner W.;
RT "New potential cell wall glucanases of Saccharomyces cerevisiae and
their involvement in mating."
RL J. Bacteriol. 180:5030-5037(1998).
CC -1- FUNCTION: Involved in cell separation.
CC -1- INTERACTION:
CC P43603:YFR024C; NdeExp=1; IntAct=EBI-16773, EBI-22380;
CC P33793:YHR016C; NdeExp=1; IntAct=EBI-16773, EBI-24460;
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- SIMILARITY: Belongs to the glycosyl hydrolase 17 family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

```

```

CC -----
DR EMBL: Z72550; CAA96729.1; -, Genomic DNA.
DR PIR: S64030; S64030.
DR INFACT: PS3189; -.
DR GeneOntology: 141076; -.
DR Ensembl: YGL028C; Saccharomyces cerevisiae.
DR SGD: S000002396; SCW11.
DR GO: GO:0009277; C:cell wall (sensu Fungi); IDA.
DR GO: GO:0007109; P:cytokinesis, completion of separation; IMP.
DR InterPro: IPR000490; Glyco_hydro_17.
DR PROSITE: PS00587; GLYCOSYL_HYDROL_F1; 1.
KM Cell wall; Complete proteome; Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 542 Probable family 17 glucosidase SCW11.
FT COMBINS 31 303 Ser/Thr-rich.
FT ACT_SITE 478 478 Nucleophile (by similarity).
FT ACT_SITE 532 532 Proton donor (by similarity).
SQ SEQUENCE 542 AA; 56447 MW; 560CB07CA1669959 CRC64;

Query Match 9.3%; Score 120.5; DB 1; Length 542;
Best Local Similarity 25.4%; Pred. No. 0.54;
Matches 51; Conservative 27; Mismatches 60; Indels 63; Gaps 8;

QY 50 NTPNASTVSAATSLVNSTNSGVAQKQNSSTSPARKSTNSLQHQQAATTS----- 103
DB 217 NTPNASTVSAATSLVNSTNSGVAQKQNSSTSPARKSTNSLQHQQAATTS----- 103
QY 104 -----SSQSKRIYPIFHYGKVDLEIRVNSLQOVKSYGVYGCIG-----ETANAF 149
DB 274 SAATTSPPKAAIAYSPYNDGCKSADAVSDLTIRKSGIKIRYVGTDCNSFTVQPA 333
QY 150 WVIINTTINDG-----STPMVVDGIFHLQNLNGVYQPDSTAETI-YANTNSGRTPTDLN 203
DB 334 VKLGIKI-NQGLHYTSSGVDSD-----DSVTTLLIQYQTNQ----- 369
QY 204 PGVSWTTNLVDPMPDFTYGH 224
DB 370 -----MDVDFPFTVGN 380

RESULT 10
Q8KB32 CHLTP PRELIMINARY; PRT; 170 AA.
AC Q8KB32;
ID Q8KB32 CHLTP PRELIMINARY; PRT; 170 AA.
DT 01-OCT-2002 (TRENBLREL. 22, Created)
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DE Hypothetical protein.
GN OrderedLocNames=CT1966;
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
CC Chlorobaculum.
OK NCBI_TaxID=1097;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey J.E., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA Radu D., Vamathaven U.J., Kouti H.M., White O., Gruber T.M.,
RA Ketchum K.A., Venter J.C., Tetteijn H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
photosynthetic, anaerobic, green-sulfur bacterium."
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL: AA006470; AAM73184.1; -, Genomic DNA.
DR TIGR: CT1966; -.
SQ Complete proteome; Hypothetical protein.
KM SEQUENCE 170 AA; 18594 MW; 7BC1EB8EA62CD4FD CRC64;

Query Match 9.3%; Score 119.5; DB 2; Length 170;

```

DR Pfam; PF07554; FIVAR; 44.

DR Pfam; PF07554; FIVAR; 44.  
DR Pfam; PF07564; DUF1542; 8







**This Page Blank (uspto)**



Sequence 19725, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 19725  
 LENGTH: 417  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-19725

Query Match 8.3%; Score 106.5; DB 2; Length 417;  
 Best Local Similarity 25.6%; Pred. No. 0.032; Mismatches 57; Indels 57; Gaps 12;  
 Matches 58; Conservative 33; Mismatches 79; Indels 57; Gaps 12;

QY 32 GSLMALIVPANTVYMRGTPRNASTVSA-----TTSLVNS----- 67  
 DB 21 GIKMLTVMNTIASL---NTGRNLNNSASLNTSLQRLSTGSRINSAKODAGLQIANRL 77  
 QY 68 ---TSSQVAKQEQNSSTPA-----HKSTNSLOHQAQ---QAAITSSQSKRIYIPF 114  
 DB 78 TSQVGNLAVATKANDGISLAQTARQALQOQSTNIIQRMEDLSLQASNGSNDSEBTAL-- 135  
 QY 115 HTYQAGDLEIRVNSLQOVKSVGVDIGETANGAFVINITIRNDSTPMEVVD-GIFHL 173  
 DB 136 --NGEVKQLQ---KELDRISNTTTGGRGLDLSGFEVASFQV---GSAANEIISVGDIM 187  
 QY 174 --QINGNVYQPDSTAETIYANTNSGTIPTDLNPGVS--MTTNLVFDM 216  
 DB 188 SAEISLNGTYFKADGGAVTAATASGTV--DIAGITGSAVAVKQDM 232

RESULT 3  
 US-10-172-502-4  
 Sequence 4, Application US/10172502  
 Patent No. 6841154  
 GENERAL INFORMATION:  
 APPLICANT: POSTER, Timothy et al.  
 TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
 FILE REFERENCE: P07263US01/BAS  
 CURRENT APPLICATION NUMBER: US/10/172,502  
 CURRENT FILING DATE: 2002-06-17  
 PRIOR APPLICATION NUMBER: US 60/298,098  
 PRIOR FILING DATE: 2001-06-15  
 NUMBER OF SEQ ID NOS: 29  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 4  
 LENGTH: 2283  
 TYPE: PRT  
 ORGANISM: Staphylococcus epidermidis  
 US-10-172-502-4

Query Match 7.9%; Score 102; DB 2; Length 2283;  
 Best Local Similarity 25.0%; Pred. No. 1.2; Mismatches 78; Indels 72; Gaps 9;  
 Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVPANTVYMRGTPRNASTVSAITSLVNST---NSSQVAKQEQNSSTPA 85  
 DB 101 FAASDAPLTSRLNTQSRIVGN--QNSTIEASTSTADSTSVTKNSSV--QTSNSDVTSS 156  
 QY 86 HKS-----TNSLOHQAQATSSQSKRIYIPFTTYGKVDLEIRVNSLQOVKSVGVD 139  
 DB 157 EKSEKVTSTNSTNSQOEKLTSTSESTSK-----NTTSSSDTKSVA-- 198

QY 140 GIGETANGAFVINITIRNDSTPMEVVDGIFHLQINGNVYQPDSTAETIYANTNSGTIP 199  
 DB 199 -----STSTEQPI-----NTSTNGSTAS--NTNSQSTTP 226  
 QY 200 TDLNPGVSMITNLVFDMPDFTYGHVQGHYSLVASMGFPGSDETTYAL 247  
 DB 227 SSVNLNKTSTSTSTAPVKLRTFSRL-----AMSTFASAAITTA 266

RESULT 4  
 US-10-172-502-18  
 Sequence 18, Application US/10172502  
 Patent No. 6841154  
 GENERAL INFORMATION:  
 APPLICANT: POSTER, Timothy et al.  
 TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
 FILE REFERENCE: P07263US01/BAS  
 CURRENT APPLICATION NUMBER: US/10/172,502  
 CURRENT FILING DATE: 2002-06-17  
 PRIOR APPLICATION NUMBER: US 60/298,098  
 PRIOR FILING DATE: 2001-06-15  
 NUMBER OF SEQ ID NOS: 29  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 18  
 LENGTH: 485  
 TYPE: PRT  
 ORGANISM: Staphylococcus epidermidis  
 US-10-172-502-18

Query Match 7.7%; Score 99; DB 2; Length 485;  
 Best Local Similarity 24.5%; Pred. No. 0.23; Mismatches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 42 NTVMIRGTPRNASTVSAITSLVNST---NSSQVAKQEQNSSTPAHKS-----TNS 91  
 DB 11 NTQSETVGN--QNSTIEASTSTADSTSVTKNSSV--QTSNSDVTSSQSKRIYIPFT 66  
 QY 92 LOHAQQAATSSQSKRIYIPFTTYGKVDLEIRVNSLQOVKSVGVDIGETANGAFV 151  
 DB 67 TSNQOEKLTSTSESTSK-----NTTSSSDTKSVA----- 96  
 QY 152 INITIRNDSTPMEVVDGIFHLQINGNVYQPDSTAETIYANTNSGTIPTDLNPGVSMITN 211  
 DB 97 -----STSTEQPI-----NTSTNGSTAS--NTNSQSTTPSSVNLNKTSTST 136  
 QY 212 LVFDMPDFTYGHVQGHYSLVASMGFPGSDETTYAL 247  
 DB 137 TSTAPVKLRTFSRL-----AMSTFASAAITTA 164

RESULT 5  
 US-09-248-796A-15273  
 Sequence 15273, Application US/09248796A  
 Patent No. 6747137  
 GENERAL INFORMATION:  
 APPLICANT: Keith Weinstein et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 FILE REFERENCE: 107196.132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 15273  
 LENGTH: 994  
 TYPE: PRT  
 ORGANISM: Candida albicans  
 US-09-248-796A-15273

Query Match 7.5%; Score 97; DB 2; Length 994;  
 Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

Best Local Similarity 22.3%; Pred. No. 1.1;  
Matches 55; Conservative 34; Mismatches 82; Indels 76; Gaps 9;

QY 42 NNVVIRKGTTPRNASTVSAVTSIV--NSTNSSQVAKQONS-----TSPAKST 89  
DB 147 STINOLASTJNSTLTITPTPTSLVPPFSYTNSSRVSTWRNSSBSMWLEKSSQLLS 206  
QY 90 NSLOHQAHOAATSSSQ-----SKLAIPIPHYGVKXGDLKIRNSIQVKSVCYDDI 141  
DB 207 TSPFNSSSTASTTESSELASATTSOSLNHSSSSSVLSSSEADSSSESVETGSS 266  
QY 142 GETANGAFWVNIITIRNDSTPEVVDGIFHLQNLNGVYOPDST--ARIYANTSGTI 198  
DB 267 DETAS-----NYSGDLFKALDTNAPPTVARSF--I 295  
QY 199 PFDLNPV-----SMTNLVFPMDPMTYTG-HVGOHYSLVASMGF----- 237  
DB 236 PLTIIPAGVNNKPGICTNKFYTNLLGNQDFWVYPLPYGLYMSKTSYGFVAVQHNNVSDR 355  
QY 238 -FGSDT 243  
DB 356 VFGSINT 362

## RESULT 6

US-09-228-986-72  
; Sequence 72, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 72  
; LENGTH: 910  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-72

Query Match 7.5%; Score 96.5; DB 2; Length 910;  
Best Local Similarity 22.6%; Pred. No. 1.1;  
Matches 58; Conservative 40; Mismatches 98; Indels 61; Gaps 15;

QY 10 ILVPIYMLF---IQGRNRLR--FAGSLMALIVFANTVMIRGTPRNASTVSATTS 63  
DB 302 IALPYNQSLGIVPMELGKLTQLRPFQGN-----YFISGNTWR---CPILTA 346  
QY 64 LVNSTNSSQVAKQONS-----TSPAKSTNSLOH---AQHOAATSSSQSKLRYPH 115  
DB 347 LSNCSDLQYVDSENNFTQQLPFSIGHLS-KKLYHLDGSLNLAGIPIPAIGNLSLFTL 405  
QY 116 TYGK--VGDLKIRNSIQVKSVCYDGIGETANGAFWVNIITIRNDSTPEVVD--- 168  
DB 406 NIGRNYFTGSIPISSLIIMLQKRLTYMD-----SNNL-----QGNIPMEIGQLKSL 450  
QY 169 GIFHL--QNLNGVYOPDSTARI---YANTSGTIPTDLPNGVSMTTN-LVPDMPDFMT 221  
DB 451 GLVYLSGNLSGKI--PDFVANLQQLRYLYLNHNLGSDINNLKCVALLDLDSYNTL 508  
QY 222 YGHVGOHYSLVASMGF 238  
DB 509 SGHIFQELAGLANIAFY 525

## RESULT 7

US-10-101-464A-72  
; Sequence 72, Application US/10101464A  
; Patent No. 6768044

## GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 72  
; LENGTH: 910  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-10-101-464A-72

Query Match 7.5%; Score 96.5; DB 2; Length 910;  
Best Local Similarity 22.6%; Pred. No. 1.1;  
Matches 58; Conservative 40; Mismatches 98; Indels 61; Gaps 15;

QY 10 ILVPIYMLF---IQGRNRLR--FAGSLMALIVFANTVMIRGTPRNASTVSATTS 63  
DB 302 IALPYNQSLGIVPMELGKLTQLRPFQGN-----YFISGNTWR---CPILTA 346  
QY 64 LVNSTNSSQVAKQONS-----TSPAKSTNSLOH---AQHOAATSSSQSKLRYPH 115  
DB 347 LSNCSDLQYVDSENNFTQQLPFSIGHLS-KKLYHLDGSLNLAGIPIPAIGNLSLFTL 405  
QY 116 TYGK--VGDLKIRNSIQVKSVCYDGIGETANGAFWVNIITIRNDSTPEVVD--- 168  
DB 406 NIGRNYFTGSIPISSLIIMLQKRLTYMD-----SNNL-----QGNIPMEIGQLKSL 450  
QY 169 GIFHL--QNLNGVYOPDSTARI---YANTSGTIPTDLPNGVSMTTN-LVPDMPDFMT 221  
DB 451 GLVYLSGNLSGKI--PDFVANLQQLRYLYLNHNLGSDINNLKCVALLDLDSYNTL 508  
QY 222 YGHVGOHYSLVASMGF 238  
DB 509 SGHIFQELAGLANIAFY 525

## RESULT 8

US-08-883-534-6  
; Sequence 6, Application US/08883534  
; Patent No. 5846777  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

RESULT 10  
 US-09-107-532A-6217  
 ; Sequence 6217 Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LYNN A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Waltham  
 ; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD-ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Atinello, Pamela Deneke

```

;
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6217:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 522 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHEetical: YES
;   ORIGINAL SOURCE:
;     ORGANISM: Enterococcus faecium
;   FEATURE:
;     NAME/KEY: misc_feature
;     LOCATION: (8) LOCATION 1...522
;     SEQUENCE DESCRIPTION: SEQ ID NO: 6217:
US-09-107-532A-6217

Query Match      7.3%; Score 94.5; DB 2; Length 522;
Best Local Similarity 20.2%; Pred. No. 0.75;
Matches 50; Conservative 41; Mismatches 96; Indels 61; Gaps 11;

OY 37 LIVFANTVYMTKGNTPRNASVTSATSL-VNSTSSQVAKQKQNSSTSPAKHSTNSLQHA 95
DB 211 LILAKETIYGYTPOPKASHSHINAASHMTNSTNDTHLOKTEGLADPKQKVNLTQKI 270
OY 96 Q-----HOAAATSSSQSKLRYIPFHT-YGKVGDEIRVNSLQOVKSVGY--DGIGET 144
DB 271 NDKLDTETHSLVMTTDSGAMVWVPFIDSOHWRBETILTGLIRE---GYMDNTYDI 326
OY 145 ANGAFWVIN-----ITIRNDGSTPMEVY-----DGIFHLQNLNG 178
DB 327 SADPSWFIYSSGDDVRLITSRDNGKTMQTSITIGSVQPIRYKAVFYGNQDLFICTASG 386
OY 179 NVYQPDSTAEIYANTNSGTIPDLNPGVSMTNLVFMDPDMFTYGHVQGHYSLV-ASNGF 237
DB 387 ---MSAEIYLYTTWDGI--TWIKSGSTMINQPV-----QNVSPVYTSIGF 428

OY 238 FGSDETV 245
DB 429 VSTRSLP 436

RESULT 11
US-09-009-620-2
; Sequence 2, Application US/09009620A
; Patent No. 6127603
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357C
; CURRENT APPLICATION NUMBER: US/09/009,620A
; CURRENT FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-009-620-2

Query Match      7.3%; Score 94; DB 2; Length 349;
Best Local Similarity 21.3%; Pred. No. 0.45;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;
```

```

OY 102 TSSSQSKL-----RYIPFH-----TYKVGDEIRVNSLQOVKSVGYDGIQTANGAF 149
DB 127 TSKAVNRLTANGKEVPANNNTAHPEKTVYDKIYPIPKIKGLKQPSLSQDDIALGN--- 183
OY 150 WVINI-----IRNDGSTPMEVVDIPIHLQNLNGVYQPDSTAEIYANT---NSGTI 198
DB 184 -VKNIKRVNGKYYTYKEDGTIQKN-----YALNINNGKTFPFDGTALSNNTLPSKGN 236
OY 199 PTDLNPGVSMTNLVFMDPDMFTYGHVQGHY 229
DB 237 TNNNTNSFAQYQNVYS--TDVANPEHY-DHY 265

RESULT 12
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match      7.3%; Score 94; DB 2; Length 1375;
Best Local Similarity 21.3%; Pred. No. 3.7;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;

OY 51 TPRNASVTSATSLVNSTNSQVAKQKQNSSTSPAKHSTNSLQHAHQHQAAT 101
DB 67 TATDTSTATSATSGPPTATVTONVSTNGSTVNTANTANFVVKPTTTSQAATTNDSK 126
OY 102 TSSSQSKL-----RYIPFH-----TYKVGDEIRVNSLQOVKSVGYDGIQTANGAF 149
DB 127 TSKAVNRLTANGKEVPANNNTAHPEKTVYDKIYPIPKIKGLKQPSLSQDDIALGN--- 183
OY 150 WVINI-----IRNDGSTPMEVVDIPIHLQNLNGVYQPDSTAEIYANT---NSGTI 198
DB 184 -VKNIKRVNGKYYTYKEDGTIQKN-----YALNINNGKTFPFDGTALSNNTLPSKGN 236
OY 199 PTDLNPGVSMTNLVFMDPDMFTYGHVQGHY 229
DB 237 TNNNTNSFAQYQNVYS--TDVANPEHY-DHY 265

RESULT 13
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
```

```
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match 7.3%; Score 94; DB 2; Length 1375;
Best Local Similarity 21.3%; Pred. No. 3.7;
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;
```

```
QY 51 TPRNASTVASTVATSLVNSTNSQVAKQKONSSTS-----PAKSTNSLOHQAHOQAT 101
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 67 TATPTSTATSATSPPTATVTDNVSTNQSTWNTANTANFVYKPTTTSQAKTDNSDKIIT 126
QY 102 TSSSQSKL-----RYIPFH-----TYGKVDLEIRVNSLQOVKSVGYDGIGETANGAF 149
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 127 TSKAVNRLTATGKVPANNNTAHPKTYTDKIVPKPKIGKLPSSLSQDDIALGN--- 183
QY 150 WVINI#-----IRNDGTPMEVVDGIFHLQNLNGVYQPDSTAETIYAT---NSGTI 198
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 184 -VKNIRKVNKGYKYYKEDGTLOKN-----YALNINGTFFPDGTALSNNTLSKKGNI 236
QY 199 PTDLNPGVSMTTNLVFPMDPFTYGVHQHY 229
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 237 TNNNDNTNSFAQYNQVYS-TDVANFEHV-DHY 265
```

```
RESULT 14
US-09-270-767-31859
Sequence 31859, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 31859
LENGTH: 313
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-31859
```

```
Query Match 7.3%; Score 93.5; DB 2; Length 313;
Best Local Similarity 25.7%; Pred. No. 0.43;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;
```

```
QY 46 MIRNGTPRNASTVATSL-----VNSTNSQVAKQ---EONSTSPAKHSTNSLOH 94
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 142 VVSTNTSLTNTTASTSTASIMHRTKTDISHNSSGAGSQFYCELNSTSTSRHNSLDVDH 201
QY 95 AQHQA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSVGYDGIGETANGAFWVIN 153
```

```
Db 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTTHSQSWG-NGLKSHGNGT----- 249
QY 154 ITRNDGTPMEVVDGIFHLQNLNGVYQPD 185
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 250 -----TNGGSP-----HINAQLVOPSS 266
```

```
RESULT 15
US-09-270-767-47076
Sequence 47076, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47076
LENGTH: 313
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-47076
```

```
Query Match 7.3%; Score 93.5; DB 2; Length 313;
Best Local Similarity 25.7%; Pred. No. 0.43;
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;
```

```
QY 46 MIRNGTPRNASTVATSL-----VNSTNSQVAKQ---EONSTSPAKHSTNSLOH 94
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 142 VVSTNTSLTNTTASTSTASIMHRTKTDISHNSSGAGSQFYCELNSTSTSRHNSLDVDH 201
QY 95 AQHQA-ATTSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSVGYDGIGETANGAFWVIN 153
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 202 HQHASLMSASSSVSNLSLDSNSGGR-----QSGKLTTHSQSWG-NGLKSHGNGT----- 249
QY 154 ITRNDGTPMEVVDGIFHLQNLNGVYQPD 185
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 250 -----TNGGSP-----HINAQLVOPSS 266
```

```
Search completed: Apr 11 7, 2006, 16:25:19
Job time : 37.5187 secs
```



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:40:31 ; Search time 121.002 Seconds

(without alignments)  
856.362 Million cell updates/sec

Title: US-10-784-592-43

Perfect score: 1289  
Sequence: 1 KRIMVGLMIVPYIMLFIO.....YSLVSMGFRSGDITVALP 248

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA Main:

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1289	100.0	248	US-10-784-592-43	Sequence 43, Appl
2	135	10.5	331	US-09-815-242-5912	Sequence 5912, Ap
3	135	10.5	331	US-09-815-242-12872	Sequence 12872, A
4	135	10.5	331	US-09-815-242-13146	Sequence 13146, A
5	135	10.5	331	US-10-282-122A-44046	Sequence 44046, A
6	113.5	8.8	6713	US-10-282-122A-43811	Sequence 43811, A
7	107.5	8.3	2086	US-09-815-242-5639	Sequence 5639, Ap
8	107.5	8.3	5795	US-09-815-242-12610	Sequence 12610, A
9	106	8.2	2344	US-09-815-242-12713	Sequence 12713, A
10	105.5	8.2	1020	US-10-437-963-131231	Sequence 131231, A
11	105	8.1	2386	US-10-156-761-7751	Sequence 7751, Ap
12	102	7.9	2261	US-10-470-048B-60	Sequence 60, Appl
13	102	7.9	2271	US-10-282-122A-43924	Sequence 43924, A
14	102	7.9	2283	US-10-172-502-4	Sequence 4, Appl1
15	102	7.9	2283	US-11-020-509-4	Sequence 4, Appl1
16	101.5	7.9	394	US-10-125-692-14	Sequence 14, Appl1
17	101.5	7.9	394	US-10-991-347-14	Sequence 14, Appl1
18	100	7.8	982	US-11-097-143-4851	Sequence 4851, Ap
19	100	7.8	982	US-11-097-143-28995	Sequence 28995, A
20	99	7.7	194	US-10-282-122A-57709	Sequence 57709, A
21	99	7.7	485	US-11-020-509-18	Sequence 18, Appl
22	99	7.7	485	US-10-282-122A-509-18	Sequence 18, Appl
23	98.5	7.6	2398	US-10-282-122A-70176	Sequence 70176, A
24	98.5	7.6	6281	US-09-815-242-12896	Sequence 12896, A
25	98.5	7.6	10498	US-10-470-048B-440	Sequence 440, App
26	97.5	7.6	393	US-10-437-963-127450	Sequence 127450, A
27	96.5	7.5	910	US-10-101-464A-72	Sequence 72, Appl

28	96.5	7.5	910	US-10-864-252-72	Sequence 72, Appl
29	96	7.4	2233	US-10-998-197-16	Sequence 16, Appl
30	95.5	7.4	632	US-10-040-906A-4	Sequence 4, Appl1
31	95.5	7.4	632	US-11-098-545-4	Sequence 4, Appl1
32	95.5	7.4	786	US-10-369-493-16335	Sequence 16335, A
33	95.5	7.4	791	US-10-369-493-15596	Sequence 15596, A
34	95.5	7.4	805	US-10-369-493-15973	Sequence 15973, A
35	94.5	7.3	458	US-10-282-122A-53974	Sequence 53974, A
36	94.5	7.3	627	US-10-040-906A-6	Sequence 6, Appl1
37	94.5	7.3	627	US-11-098-545-6	Sequence 6, Appl1
38	94.5	7.3	632	US-11-067-557-2	Sequence 2, Appl1
39	94.5	7.3	633	US-11-067-557-32	Sequence 32, Appl1
40	94.5	7.3	633	US-11-067-557-38	Sequence 38, Appl1
41	94.5	7.3	634	US-11-067-557-166	Sequence 166, App
42	94.5	7.3	634	US-11-067-557-224	Sequence 224, App
43	94	7.3	1375	US-09-740-274-4	Sequence 4, Appl1
44	94	7.3	1375	US-10-383-930-35	Sequence 35, Appl1
45	94	7.3	1375	US-10-797-821-35	Sequence 35, Appl1

## ALIGNMENTS

RESULT 1					
US-10-784-592-43					
Sequence 43, Application US/10784592					
Publication No. US20050147983A1					
GENERAL INFORMATION:					
APPLICANT: Willing, Reinhard					
APPLICANT: Oestergaard, Peter					
APPLICANT: Lassen, Soren					
TITLE OF INVENTION: POLYPEPTIDES OF ALICYCLOBACILLUS SP.					
FILE REFERENCE: 10406.203-US					
CURRENT APPLICATION NUMBER: US/10/784,592					
NUMBER OF SEQ ID NOS: 55					
SOFTWARE: PatentIn version 3.2					
SEQ ID NO 43					
LENGTH: 248					
TYPE: PR					
ORGANISM: Alicyclobacillus sp.					
FEATURE:					
NAME/KEY: SIGNAL					
LOCATION: (1)..(41)					
FEATURE:					
NAME/KEY: mat.peptide					
LOCATION: (42)..(248)					
OTHER INFORMATION: functional polypeptide					
US-10-784-592-43					
Query Match					
Best Local Similarity 100.0%, Pred. No. 3.2e-11;					
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	KRIMVGLMIVPYIMLFIOGRNRILRFAGSLALIVPANTVMIRKNTPRNSTVGA	60		
DB	1	KRIMVGLMIVPYIMLFIOGRNRILRFAGSLALIVPANTVMIRKNTPRNSTVGA	60		
QY	61	TTSLVNSTNSOVAQOBNSSSTPAHKSNTSLQHOHOAAATSSOSKLRYPPTHYGV	120		
DB	61	TTSLVNSTNSOVAQOBNSSSTPAHKSNTSLQHOHOAAATSSOSKLRYPPTHYGV	120		
QY	121	GDLEIRVNSLQOVKSVGDIGETANGAFWVINITIRNDGSTPMVEVDGIFHLQNLNGV	180		
DB	121	GDLEIRVNSLQOVKSVGDIGETANGAFWVINITIRNDGSTPMVEVDGIFHLQNLNGV	180		
QY	181	YOPDSTAIRVANTNSGTLPTDLPVSMVTNIVFMPDMPTGYGHVGHVSLVSMGFRGS	240		
DB	181	YOPDSTAIRVANTNSGTLPTDLPVSMVTNIVFMPDMPTGYGHVGHVSLVSMGFRGS	240		
QY	241	DETTVALP 248			
DB	241	DETTVALP 248			

## RESULT 2

```
US-09-815-242-5912
; Sequence 5912, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5912
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5912
```

```
Query Match 10.5%; Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
```

```
QY 115 HTYG---KVGDLIRVNSLQGVKSYGDIETANGAFVYVITTRNDGSTPMEVVDGIF 171
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 81 HIGETVXNGDLEVTNVSVEYTKSVGSLAPTNAKGFVAVDVITKNKGKALTTIDSSMF 140
      ::||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
      ::||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 141 KLSK-GDKTFEADNTGSMNSQSDGSIENSFPLQRIINPDSTAGKIVFDVSE 192
```

## RESULT 3

```
US-09-815-242-12872
; Sequence 12872, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12872
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12872
```

```
Query Match 10.5%; Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00072;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
```

```
QY 115 HTYG---KVGDLIRVNSLQGVKSYGDIETANGAFVYVITTRNDGSTPMEVVDGIF 171
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 81 HIGETVXNGDLEVTNVSVEYTKSVGSLAPTNAKGFVAVDVITKNKGKALTTIDSSMF 140
      ::||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 172 HLQNLNGVYQPDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 218
      ::||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 141 KLSK-GDKTFEADNTGSMNSQSDGSIENSFPLQRIINPDSTAGKIVFDVSE 192
```

## RESULT 4

```
US-09-815-242-13146
; Sequence 13146, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13146
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
```

US-09-815-242-13146

Query Match 10.5%; Score 135; DB 3; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.00072;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG--KYGDLEIRVNSIQOVSVGYDGIGETANGAFVINITTRNDSTMEVVDGIF 171  
DB 81 HXIGETVKNKGDLEIVTVNSVETKMSVGPLAPTRNAKIFVADVITIKKKEKALTTIDSSMF 140

QY 172 HLQNLNGVYQPDSTABIYAN-TNSGTIPTD-----LNPVSMTTNLVFDMPD 218  
DB 141 KUKS-GDKTFEADNTGSMNSQNSDNGSINSFPLQRIINDSTAGQKIVDVSE 192

RESULT 5  
US-10-282-122A-44046  
Sequence 44046, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO 44046

LENGTH: 331

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-10-282-122A-44046

Query Match 10.5%; Score 135; DB 4; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.00072;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 115 HTYG--KYGDLEIRVNSIQOVSVGYDGIGETANGAFVINITTRNDSTMEVVDGIF 171  
DB 81 HXIGETVKNKGDLEIVTVNSVETKMSVGPLAPTRNAKIFVADVITIKKKEKALTTIDSSMF 140

QY 172 HLQNLNGVYQPDSTABIYAN-TNSGTIPTD-----LNPVSMTTNLVFDMPD 218  
DB 141 KUKS-GDKTFEADNTGSMNSQNSDNGSINSFPLQRIINDSTAGQKIVDVSE 192

RESULT 6  
US-10-282-122A-43811  
Sequence 43811, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zykkind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.

APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: Patentin version 3.1

SEQ ID NO 43811

LENGTH: 6713

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-10-282-122A-43811

Query Match 8.8%; Score 113.5; DB 4; Length 6713;  
Best Local Similarity 24.4%; Pred. No. 4.8;  
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;

QY 42 NTVMIRKNTPRMNST--VSATISLVNSTNSQVAKQKQNSSTRAHKSTNSLQHAQHA 99  
DB 899 NQNYLDADESKRNATYQVTAABEGILNKQGTGNTSKAVDUALNVTAKALNGAENLR 958

QY 100 ATSSSQSKLRIYPRHYGKVDLEIRVNSIQOVSVGYDGI--GETANGAFVINITTI 156  
DB 959 NAKTSATNTINGLPVLTQLODNLKHQVEQANV--VEVNGVKDKGNTLNTMGLAKRSI 1016

QY 157 RNDSTMEVVDGIFHLQNLNGVYQPDSTABIYANTNSGTIPTDLNPGVSMTTNLVFDMP 216  
DB 1017 QNDNTT--KTSQNYLDAADSXKNKY--NTA---VNNANGVINATNPD--NMDANALNDM 1066

RESULT 7

```
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 5639
; LENGTH: 2086
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-5639

Query Match      8.3%; Score 107.5; DB 3; Length 2086;
Best Local Similarity 24.2%; Pred. No. 3.4;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

Qy      42 NTVVIRGNTPRNAST--VSATTSILVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQHOA 99
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1323 NQNYIDADESKRNATQAVTAABGILNKGTGNTSKADVDAALNTVTRAKAALNGAENLR 1382
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      100 ATTSSQSKLRVYIPFTYTGKVDLEIRVNSLQOVKSV-GYDGIETANGAFVINITIRN 158
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1383 NTKTSATNTINGLPULTQKQDNLKHQVEQAQNVAGVGVKDKGNTLTAMGALRTSIGN 1442
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      159 DGSITPEVVDGIFHLQNLNGVYQDPSTAEIYANTNSGTFPLDLPVGSMTTNLVFDM 216
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1443 DMTT--KTSQNYVLADSDSKNNY---NTA---VNNANGVINTVNNP--NMDANALNGM 1490
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
US-09-815-242-12610
; Sequence 12610, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12610
; LENGTH: 5795
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
US-09-815-242-12610

Query Match      8.3%; Score 107.5; DB 3; Length 5795;
Best Local Similarity 24.2%; Pred. No. 14;
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

Qy      42 NTVVIRGNTPRNAST--VSATTSILVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQHOA 99
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      3600 NQNYIDADESKRNATQAVTAABGILNKGTGNTSKADVDAALNTVTRAKAALNGAENLR 3659
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      100 ATTSSQSKLRVYIPFTYTGKVDLEIRVNSLQOVKSV-GYDGIETANGAFVINITIRN 158
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      3660 NTKTSATNTINGLPULTQKQDNLKHQVEQAQNVAGVGVKDKGNTLTAMGALRTSIGN 3719
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      159 DGSITPEVVDGIFHLQNLNGVYQDPSTAEIYANTNSGTFPLDLPVGSMTTNLVFDM 216
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      3720 DMTT--KTSQNYVLADSDSKNNY---NTA---VNNANGVINTVNNP--NMDANALNGM 3767
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-09-815-242-12713
; Sequence 12713, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

QY 200 TDLNPGVSM TTNLVFDMDFMTY 222

Quality match	7.3%	Score 102;	DB 5;	Length 2261;
Best Local Similarity	25.0%	Pred. 102;		
Matches 57; Conservative	21;	Mismatches	78;	Indels 72;
		Gaps	9;	

QY 30 FAGSIMALIVPANTYVIRGNTPRNASTVATSLVNST---NSQVAKQONSSTSPA 85  
DB 89 FAASDAPLITSELTNOSTEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 144  
QY 86 HKS-----TNSLQAHQOAAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVYSGVD 139  
DB 145 EKSEKVTSTNSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 186  
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199  
DB 187 -----STSTEQPI-----NTSTNOSTAS--NNTSOSSTTP 214  
QY 200 TDLPNGVSMITNLVFDMPDFMTYGVGHVSLVASMFGFGSDETTYAL 247  
DB 215 SSVNINKTSTSTSTAPVKLTFRSL-----AMSTFASAAATTAV 254

## RESULT 13

US-10-282-122A-43924  
; Sequence 43924, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining prior application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 43924  
LENGTH: 2271  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-10-282-122A-43924

Query Match 7.9%; Score 102; DB 4; Length 2271;  
Best Local Similarity 25.0%; Pred. No. 12;  
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;  
QY 30 FAGSIMALIVPANTYVIRGNTPRNASTVATSLVNST---NSQVAKQONSSTSPA 85

DB 89 FAASDAPLITSELTNOSTEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 144  
QY 86 HKS-----TNSLQAHQOAAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVYSGVD 139  
DB 145 EKSEKVTSTNSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 186  
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199  
DB 187 -----STSTEQPI-----NTSTNOSTAS--NNTSOSSTTP 214  
QY 200 TDLPNGVSMITNLVFDMPDFMTYGVGHVSLVASMFGFGSDETTYAL 247  
DB 215 SSVNINKTSTSTSTAPVKLTFRSL-----AMSTFASAAATTAV 254

## RESULT 14

US-10-172-502-4  
; Sequence 4, Application US/10172502  
; Publication No. US20030185833A1

## GENERAL INFORMATION:

APPLICANT: FOSTER, Timothy et al.  
TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
FILE REFERENCE: P07263US01/BAS  
CURRENT APPLICATION NUMBER: US/10/172,502  
CURRENT FILING DATE: 2002-06-17  
PRIOR APPLICATION NUMBER: US 60/298,098  
PRIOR FILING DATE: 2001-06-15  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 2283  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-10-172-502-4

Query Match 7.9%; Score 102; DB 4; Length 2283;  
Best Local Similarity 25.0%; Pred. No. 12;  
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSIMALIVPANTYVIRGNTPRNASTVATSLVNST---NSQVAKQONSSTSPA 85  
DB 101 FAASDAPLITSELTNOSTEVGN--QNSTTIEASTSTADSTVTKNSSSV--QTSNSDVTSS 156  
QY 86 HKS-----TNSLQAHQOAAATTSSQSKLARIIPHTYGVKXGDLIRVNSLQOVYSGVD 139  
DB 157 EKSEKVTSTNSTNSQOEKLTSTSESTSSK-----NTTSSSDTKSVA-- 198  
QY 140 GIGETANGAFVINITINDGSTPMEVVDGIFHLQNLNGVYQPDSTAIYANTNSGTIP 199  
DB 199 -----STSTEQPI-----NTSTNOSTAS--NNTSOSSTTP 226  
QY 200 TDLPNGVSMITNLVFDMPDFMTYGVGHVSLVASMFGFGSDETTYAL 247  
DB 227 SSVNINKTSTSTSTAPVKLTFRSL-----AMSTFASAAATTAV 266

## RESULT 15

US-11-020-509-4  
; Sequence 4, Application US/11020509  
; Publication No. US20050106648A1

## GENERAL INFORMATION:

APPLICANT: FOSTER, Timothy et al.  
TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
FILE REFERENCE: P07263US02/BAS  
CURRENT APPLICATION NUMBER: US/11/020,509  
CURRENT FILING DATE: 2004-12-27  
PRIOR APPLICATION NUMBER: US 10/172,502  
PRIOR FILING DATE: 2002-06-17  
PRIOR APPLICATION NUMBER: US 60/298,098  
PRIOR FILING DATE: 2001-06-15  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4  
; LENGTH: 2283  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-11-020-509-4

Query Match 7.9%; Score 102; DB 6; Length 2283;  
Best Local Similarity 25.0%; Pred. No. 12;  
Matches 57; Conservative 21; Mismatches 78; Indels 72; Gaps 9;

QY 30 FAGSLMALIVFANTVTMIRGNTPRNASTVSATTSLVNST---NSSQVAKQKQNSSSTSPA 85  
DB 101 FAASDBAPLITSEILNTQSEIVGN--QNSTTIEASTADSTSVTKNSSSV--QTSNSDPTVSS 156  
QY 86 HKS-----TNSIQHQHQAAATSSSQSKLRIIPHTYGVKVDLEIRVNSLQOVKSYGYD 139  
DB 157 EKSEKVTSTNSTNSQOEKLTSTSESTSK-----NTTSSSDPTKSYA-- 198  
QY 140 GIGETANGAFWVINITIRNDGSTPMEVVDGIFHLQNLNGNVYQPDSTAETIYANTNSGTIP 199  
DB 199 -----STSTEQPI-----NTSTNQSTAS--NNTSQSTTP 226  
QY 200 TDINPGVSMITTNLVFPDMPFTMYGHVGOHYSLVASMGFFGSDETTYAL 247  
DB 227 SSVNLNKTSTTSTSTAPVKLRTPSRL-----AMSTFASAAATTAY 266

Search completed: April 7, 2006, 16:44:39  
Job time : 123.002 secs

**This Page Blank (uspto)**









;; PRIOR FILING DATE: 1998-07-28  
;; NUMBER OF SEQ ID NOS: 24  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 4  
;; LENGTH: 729  
;; TYPE: PRT  
;; ORGANISM: Mus sp.  
;; FEATURE:  
;; OTHER INFORMATION: mouse taste cell polypeptide (TCP) #2 amino acid  
;; OTHER INFORMATION: sequence  
US-11-060-920-4

Query Match 6.8%; Score 89.5; DB 7; Length 729;  
Best Local Similarity 28.5%; Pred. No. 3.1;  
Matches 37; Conservative 14; Mismatches 44; Indels 35; Gaps 6;

QY 42 NTVMIRGNTPRNAs-----TTSATSLVNSTN-SSQVAKQKQNS-----TSPAKSTNS 91  
DB 243 NTVALDTGPRVSGADSDSPQTPSTJSDSFKTSNLPQIALQSHQGLMPTSPHMPPLS 302  
QY 92 LQH-----AQHQAATSSS-----QSLARVYIPHTYK--VGDLEIR 126  
DB 303 LQHSPPSTHSSGFTSSVHADPTLASTLPHPQDMSLQDLSFGSRHTHTSVPR 362  
QY 127 VNSLQOVXSV 136  
DB 363 INSNRFTKAV 372

RESULT 9  
US-11-212-443-111  
;; Sequence 111, Application US/11212443  
;; Publication No. US20050287165A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Scalco, Enzo  
;; APPLICANT: Massignani, Vega  
;; APPLICANT: Rappucci, Rino  
;; APPLICANT: Pizza, Mariagrazia  
;; APPLICANT: Grandi, Guido  
;; TITLE OF INVENTION: Meningococcal Antigens  
;; FILE REFERENCE: CHIR0159  
;; CURRENT APPLICATION NUMBER: US/11/212,443  
;; CURRENT FILING DATE: 2005-08-24  
;; PRIOR APPLICATION NUMBER: US/09/302,626  
;; PRIOR FILING DATE: 1999-04-30  
;; PRIOR APPLICATION NUMBER: PCT/IB99/00103  
;; PRIOR FILING DATE: 1999-01-14  
;; NUMBER OF SEQ ID NOS: 195  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 111  
;; LENGTH: 600  
;; TYPE: PRT  
;; ORGANISM: Neisseria meningitidis  
US-11-212-443-111

Query Match 6.8%; Score 87.5; DB 7; Length 600;  
Best Local Similarity 20.4%; Pred. No. 3.6;  
Matches 46; Conservative 42; Mismatches 120; Indels 17; Gaps 8;

QY 24 MRRILRFGSLMALIVFANTYV--MIRGNTPRNASTVSAT--TSLVNSTNSQVAKQKQNS 79  
DB 1 MNRKIRI--TWNLALNAVVSSELTNRHTKASATVAVALATLISATVQANATDEEDN 57  
QY 80 SSTSPAHKSTNSLQAHQAAT-----TSSQSKLRVYIPHTYKGVGDLEIRVNSLQOVK 134  
DB 58 EDLEEVVRAVPLVSHSDKEGKEKEVGASNLVYYPDKNVLKAGTTLLAAGNLKIK 117  
QY 135 SVGYDGIETANGAFVINITIRNDGSTPEVVDGIFHLQNLNGVNYQPDSTAEIYANTN 194  
DB 118 QNTNMENTNMASTFTYSLKDLTGLINVTETKSLFGANGKVAIIS-DTGLANPAKET 176  
QY 195 SGT--IPT-DLNP-GVSMITTNLVFDMPDFTYGHVGOHYSLVASV 235

DB 177 AGTNGDPTVHLNGIGSTLTDTLAGSSASHVDAGNOSTHYTRAASI 221

RESULT 10  
US-11-060-920-5  
;; Sequence 5, Application US/11060920  
;; Publication No. US20050244378A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Kaufman, Paul L  
;; APPLICANT: Liu, Xuyang  
;; TITLE OF INVENTION: Method for Treating Glaucoma  
;; FILE REFERENCE: 960296.00149  
;; CURRENT APPLICATION NUMBER: US/11/060,920  
;; CURRENT FILING DATE: 2005-02-18  
;; PRIOR APPLICATION NUMBER: US 60/545,723  
;; PRIOR FILING DATE: 2004-02-18  
;; NUMBER OF SEQ ID NOS: 7  
;; SOFTWARE: PatentIn version 3.2  
;; SEQ ID NO 5  
;; LENGTH: 721  
;; TYPE: PRT  
;; ORGANISM: Clostridium botulinum  
US-11-060-920-5

Query Match 6.8%; Score 87.5; DB 7; Length 721;  
Best Local Similarity 21.5%; Pred. No. 4.6;  
Matches 45; Conservative 39; Mismatches 96; Indels 29; Gaps 8;

QY 28 LRFAGSLMALIVFANTYVMIRGNTPRNASTVSATTS-----LVNSTNSQVAK--QEQ 78  
DB 312 LQLAGGIRP--VFSMSASANYSHTWONTSTVDDTTGESFQGLSINTGSAVINPNIIRY 369  
QY 79 NSTSPARK--STNSLQAHQAATSSSQSL-RYI-PFHTYKGVGDLEIRVNSLQOVK 134  
DB 370 NTRAPVNTVPTTITVIDKQSVATIKQBSLIGDLYLPGGTYPIIGRPPMALNTMDQPS 429  
QY 135 S-----VGYDGIETANGAFVINITIRNDGSTPEVVDGIFHLQNLNGVNYQPDSTAEIY 190  
DB 430 SRLPNNVNTQKSLIDNGTVALSTS-----QFGNRAKNSNGLVTDGNNMGPY 479  
QY 191 ANTNSGTIPDLPNGVSMITTNLVFDMPDPF 219  
DB 480 LGITKSTVASLTLSPSGQTTQVAVAVAPNF 508

RESULT 11  
US-11-052-554A-95  
;; Sequence 95, Application US/11052554A  
;; Publication No. US2005028866A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sachdeva, et al.  
;; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
;; FILE REFERENCE: 30853/40359A  
;; CURRENT APPLICATION NUMBER: US/11/052,554A  
;; CURRENT FILING DATE: 2005-02-07  
;; PRIOR APPLICATION NUMBER: US 60/589,227  
;; PRIOR FILING DATE: 2004-07-20  
;; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
;; PRIOR FILING DATE: 2004-02-06  
;; NUMBER OF SEQ ID NOS: 763  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 95  
;; LENGTH: 1237  
;; TYPE: PRT  
;; ORGANISM: Helicobacter pylori J99  
US-11-052-554A-95

Query Match 6.7%; Score 87; DB 7; Length 1237;  
Best Local Similarity 21.1%; Pred. No. 11;  
Matches 52; Conservative 37; Mismatches 70; Indels 88; Gaps 12;

QY 38 IVFANTVMIRGNTP---RNASTVSATSLVNSTNSQVAKQK-----QNSSTSPA 85

```

Db      459 LVATSKVQTINGEBOIGVNSFNLVGVMSVNSLKTSENLQKNAKILCNQSOSTSIC 518
      86 HKSTNSLQ---HAQHO-----AATSSSQ-----SKLRIPIPTHYGVKVDLEIRVNSLQ 131
      519 NSSGGGLSIGNAQLQNLISPTNGTITNTQAQSNASKLAM-----VWVNNBE 566
Qy      132 OYKSVGYD-----GIGETANGAFWVINITIRNDGSTPMV-----VVG 169
      567 EAKTINFNQSGPTTQSSNSVWGA---LNTVLQVNSNQOSTQSAFQNOENNIQAMANA 623
Qy      170 IFHLQNLNGVYQPDST-----AEIY--ANTNSGTTPTDNL-----P 204
      624 LVNTSNPNNGNQNLTTNNNQDLRIQLRANFQULINTINQVPTDMALINQSQVQOTS 683
Qy      205 GYSMTTN 211
      684 GSASTTN 690
Db

RESULT 12
US-11-087-099-3488
; Sequence 3488, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3488
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Clostridium acetobutylicum
US-11-087-099-3488

Query Match      6.6%; Score 85; DB 7; Length 500;
Best Local Similarity 25.1%; Pred. No. 4.8;
Matches 42; Conservative 22; Mismatches 65; Indels 38; Gaps 5;

Qy      51 TPRNST-----VSATTSLVNSTSSQVAKOBONSTSPAHKSTNSLQHAQQAATTSS 105
      139 TFSNMSNKIKMTSSGNSIVSDNSGKITAIABSGSTATA-----TTTAASS 245
Qy      106 OSKLRIPIPTHYGVKVDLEIRVNSLQOYKSVGYDGISETANGAFWVINITIRNDSTPMV 165
      246 DSQKTATCVVTVNKRTPITPKDNNL-----ETS-----IRNMINKPTGSLYKE 288
Db
Qy      166 VVDGIFHLQNLNGVYQPD-----STAEIYANTNSGTTPTDNLPGVSMT 209
      289 DVQGITSPYNTNPNIHLDGIELNLEPLTFSNTPIKDLTFLKSLT 335
Db

RESULT 13
US-10-055-877-140
; Sequence 140, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Velizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Rateili, Luca
; APPLICANT: Kékuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernusen, Bryan
; APPLICANT: Andrew, David

```

```

; APPLICANT: Mezas, Peter
; APPLICANT: Patlurajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Bisen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vermet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Caeman, Stacie
; APPLICANT: Boldog, Ferenc
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 140
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-055-877-140

Query Match      6.6%; Score 85; DB 6; Length 1765;
Best Local Similarity 24.6%; Pred. No. 27;
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;

Qy      59 SATTSLVNSTSSQVAKOBONSTSPAHKSTNSLQHAQQAATTSSSQSKLRIPIPTHY 118
      670 NNAVNLMSKSGTIFHLAAQEDRVNAVAVLVNQCNAV-----DAQTKSGITPLAVGC 721
Qy      119 KVGDLEIRVNSLQOYKSV-----GYDGISETA--NGAFWVINITIRNDGSTPMVVDGI 170
      722 HYGNIKI--VNFLQHSAAVNAKTKNGYTAHLQAQGGHTHIINVLQNNAS--FNEL----- 775
Qy      171 FHLQNLNG-----YQPDSS-----TAEIYANTNSGTTPTDNLPGVSMTTLVVD 215
      776 -----TVNGNTALAIARLGLYSVVDLTKEVTEIEIWTIT---TTEKHKNVPEYINBVL 828
Db
Qy      216 MPD 218
      829 MSD 831
Db

RESULT 14
US-10-055-877-141
; Sequence 141, Application US/10055877
; Publication No. US20050288241A1

```

GENERAL INFORMATION:  
APPLICANT: Decristofaro, Marc  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Miller, Charles  
APPLICANT: Tchenev, Velizar  
APPLICANT: Zhong, Mei  
APPLICANT: Anderson, David  
APPLICANT: Ballinger, Robert  
APPLICANT: Gerlach, Valerie  
APPLICANT: Spyrek, Kimberly  
APPLICANT: Rattelli, Luca  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Guo, Xiaojia  
APPLICANT: Zernusen, Bryan  
APPLICANT: Andrew, David  
APPLICANT: Mezed, Peter  
APPLICANT: Paturajan, Meera  
APPLICANT: Burgess, Catherine  
APPLICANT: Elsen, Andrew  
APPLICANT: Wolenc, Adam  
APPLICANT: Baumgartner, Jason  
APPLICANT: Shinkets, Richard  
APPLICANT: Gusev, Vladimir  
APPLICANT: Vermet, Corine  
APPLICANT: Taupier Jr., Raymond  
APPLICANT: Pena, Carol  
APPLICANT: Shenoy, Suresh  
APPLICANT: Li, Li  
APPLICANT: Casman, Stacie  
APPLICANT: Boldog, Ferenc  
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
FILE REFERENCE: 21402-251  
CURRENT APPLICATION NUMBER: US/10/055,877  
PRIOR FILING DATE: 2002-01-22  
PRIOR APPLICATION NUMBER: 60/262,892  
PRIOR FILING DATE: 2001-01-19  
PRIOR APPLICATION NUMBER: 60/263,598  
PRIOR FILING DATE: 2001-01-23  
PRIOR APPLICATION NUMBER: 60/263,799  
PRIOR FILING DATE: 2001-01-24  
PRIOR APPLICATION NUMBER: 60/264,117  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/264,139  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/264,478  
PRIOR FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: 60/263,351  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: 60/272,870  
PRIOR FILING DATE: 2001-03-02  
PRIOR APPLICATION NUMBER: 60/275,990  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/275,927  
PRIOR FILING DATE: 2001-03-14  
Remaining Prior Application data removed - See file wrapper or PALM.  
NUMBER OF SEQ ID NOS: 512  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 141  
LENGTH: 1940  
TYPE: PRT  
ORGANISM: Mus musculus  
US-10-055-877-141

Query Match 6.6%; Score 85; DB 6; Length 1940;  
Best Local Similarity 24.6%; Pred. No. 31;  
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;  
QY 59 SATSLVNSTNSQVAKQONSSTSPAHKSTNSLQHAQQAATTSSQSKLYIPFHTYG 118  
DB 670 NANVLSKSGSLTFLHAAQDRVAVAVLVQGAHV-----DAQTKMGTPLAVGC 721  
QY 119 KVGDLIRVNSLQGVKSV-----GYDQIGETA-NGAFWVINITIRNDGSTPMEVVDGI 170

DB 722 HYGNIKI-VNFILOHSAKNVAKTKNGYTAHQAAQGHTHIIIVLLQNNAS-PNEL----- 775  
QY 171 FHLQNLN-----YQPDSS-----TAEIYANTNSGTIPFDLPAGVSMTNLYPD 215  
DB 776 ----TVNGNTALAIARRLGYISVDTLKVTEBIMTTT---TTEKHQNNVPEETNEVLD 828  
QY 216 MPD 218  
DB 829 MSD 831  
RESULT 15  
US-11-096-568A-29105  
Sequence 29105, Application US/11096568A  
Publication No. US20060048240A1  
GENERAL INFORMATION:  
APPLICANT: Alexandrov, Nikolai et al.  
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
FILE REFERENCE: 2750-1592PUS2  
CURRENT APPLICATION NUMBER: US/11/096,568A  
CURRENT FILING DATE: 2005-04-01  
NUMBER OF SEQ ID NOS: 34471  
SEQ ID NO 29105  
LENGTH: 946  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(946)  
OTHER INFORMATION: Ceres Seq. ID no. 4263811  
US-11-096-568A-29105  
Query Match 6.6%; Score 84.5; DB 7; Length 946;  
Best Local Similarity 21.7%; Pred. No. 13;  
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;  
QY 52 PRNASTVSATSLVNSTNSQVAKQONSSTSPAHKSTNSLQHAQQAATTSSQSKLYR 111  
DB 59 PDNSVLSLSNFDLSNSSFPLVLCNQLTESLDVSNRRLSIPGCVTCERLIALKH 118  
QY 112 IPFHT-----YKQVDLIRVNSLQ-QVKSQYDQIGETANGAFWVINITI-RN 158  
DB 119 INFSTNKRSTSPGRGFSKLAVLDPESHVLSGNVGDYFGDGLVQLRS-----LNLSPNRL 173  
QY 159 DGSPTMEVVDGIFHLQNLNGVNVQPDSTAETIYANTNSGTITFDLPNGVSMKTNLVFMDP 218  
DB 174 TGSVPVHLTKSLBKL-----EVSDNSLSTIGIKDYQELT---LIDLS 216  
QY 219 PMYGHVQGHYSLVASWG 236  
DB 217 NQNLG-----SIPSSLDG 228

Search completed: April 7, 2006, 16:40:50  
Job time : 18.448 secs



Query Match 100.0%; Score 1074; DB 9; Length 207;  
Best Local Similarity 100.0%; Pred. No. 1e-89;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVVMIRGNTPRNASTVSATTSVLNSTNSGVAKOEONSTSPARKSTNSLOHQAHOAT 60  
DB 1 NTVVMIRGNTPRNASTVSATTSVLNSTNSGVAKOEONSTSPARKSTNSLOHQAHOAT 60  
QY 61 TSSSOSKLRYPFHTYGVGDLEIRVNSLQOVKSVGYDGETANGAFWVINITIRNDG 120  
DB 61 TSSSOSKLRYPFHTYGVGDLEIRVNSLQOVKSVGYDGETANGAFWVINITIRNDG 120  
QY 121 TPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTITPDLPVGSMTTNLVFMDPDMT 180  
DB 121 TPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTITPDLPVGSMTTNLVFMDPDMT 180  
QY 181 YGHVGQHYSLVASMGPFGSDETTVALP 207  
DB 181 YGHVGQHYSLVASMGPFGSDETTVALP 207

## RESULT 2

AEBA45569 standard; protein; 248 AA.

AEBA45569;

22-SEP-2005 (first entry)

Allcyclobacillus sp. functional polypeptide, SEQ ID NO: 43.

Feedstuff; food; detergent; surfactant; pulp; functional polypeptide.

Allcyclobacillus sp.; DSM 15716.

Key Location/Qualifiers

Peptide 1..41 /label= Signal\_peptide

Misc-difference 1 /note= "Encoded by GTC"

Protein /note= "Allcyclobacillus sp. mature functional polypeptide"

US2005147983-A1.

07-JUL-2005.

23-FEB-2004; 2004US-00784592.

06-JAN-2004; 2004DK-00000010.

04-FEB-2004; 2004DK-00000165.

(NOVO) NOVOZYMES AS.

Willing R, Lassen SF, Ostergaard PR;

WPI, 2005-511773/52.

N-PsDB; AEB45544.

New functional polypeptides having function and amino acid sequence similar to known specific bacterial enzymes useful in industrial, research and household applications e.g. detergents and food.

Example 2, SEQ ID NO 43; 83pp; English.

The present invention relates to functional polypeptides encoded by polynucleotides comprised in the genome of Allcyclobacillus sp. DSM 15716. The invention is useful in industrial, research and household processes such as in composition to prepare feed and food, in detergent formulations and for treating lignocellulosic fabric and pulp. The present sequence is Allcyclobacillus sp. DSM 15716 functional

CC polypeptide.

XX Sequence 248 AA;

Query Match 100.0%; Score 1074; DB 9; Length 248;  
Best Local Similarity 100.0%; Pred. No. 1.3e-89;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVVMIRGNTPRNASTVSATTSVLNSTNSGVAKOEONSTSPARKSTNSLOHQAHOAT 60  
DB 42 NTVVMIRGNTPRNASTVSATTSVLNSTNSGVAKOEONSTSPARKSTNSLOHQAHOAT 101  
QY 61 TSSSOSKLRYPFHTYGVGDLEIRVNSLQOVKSVGYDGETANGAFWVINITIRNDG 120  
DB 102 TSSSOSKLRYPFHTYGVGDLEIRVNSLQOVKSVGYDGETANGAFWVINITIRNDG 161  
QY 121 TPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTITPDLPVGSMTTNLVFMDPDMT 180  
DB 162 TPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTITPDLPVGSMTTNLVFMDPDMT 221  
QY 181 YGHVGQHYSLVASMGPFGSDETTVALP 207  
DB 222 YGHVGQHYSLVASMGPFGSDETTVALP 248

## RESULT 3

AEBA8755 standard; protein; 248 AA.

AEBA8755;

22-SEP-2005 (first entry)

Allcyclobacillus sp. DSM 15716 functional polypeptide.

Protein production; protein secretion; surfactant; feedstuff; food.

Allcyclobacillus sp.; DSM 15716.

Key Location/Qualifiers

Peptide 1..41 /label= Signal\_peptide

Misc-difference 1 /note= "Encoded by GTC"

Protein /label= Mature\_polypeptide

WO200506339-A2.

21-JUL-2005.

06-JAN-2005; 2005MO-DK000004.

06-JAN-2004; 2004DK-00000010.

04-FEB-2004; 2004DK-00000165.

23-FEB-2004; 2004US-00784592.

25-FEB-2004; 2004DK-00000298.

(NOVO) NOVOZYMES AS.

Willing R, Lassen SF, Ostergaard PR;

WPI, 2005-506869/51.

N-PsDB; AEB48730.

New mature functional polypeptide of Allcyclobacillus sp., used for commercial research purposes, specifically for cleaning a cellulosic fabric, preparing food or feed additive, and for treating lignocellulosic materials and pulp.

Claim 3, SEQ ID NO 43; 151pp; English.

The present invention is based on the finding of a strain of



CC Alicyclobacillus, namely Alicyclobacillus sp. DSM 15716, which grows at  
CC low pH (approximately 4-5) and at high temperature (50-60 degrees C). It  
CC is an object of the invention to identify and provide polypeptides CC  
CC secreted from Alicyclobacillus sp. DSM 15716 because such polypeptides  
CC may be used for industrial purposes and may also be produced in  
CC industrially relevant processes and amounts. Thus, the invention provides  
CC isolated mature functional polypeptides which are at least 90% identical  
CC to, and exhibit and same function as a corresponding secreted polypeptide  
CC ABB48736-ABB48762 from Alicyclobacillus sp. DSM 15716, and also provides  
CC polynucleotides ABB48713-ABB48737 encoding such polypeptides. The  
CC polypeptides are preferably enzymes having acid endoglucanase, acid  
CC cellulase, aspartyl protease, multi copper oxidase, serine-carboxyl  
CC protease, serine protease, H<sub>2</sub>O<sub>2</sub>-like serine protease, disulfide  
CC isomerase, gamma-D-glutamyl-L-amino acid endopeptidase, endo-beta-N-  
CC acetylglucosaminidase, peptidyl-prolyl-isomerase, acid phosphatase,  
CC phytase, phospholipase C, polysaccharide deacetylase, xylanase  
CC deacetylase or sulfite oxidase activity. A composition comprising such an  
CC enzyme can be used in a detergent composition, or a food or feed  
CC composition. Also claimed are: nucleic acid constructs, recombinant  
CC expression vectors and host cells; methods of producing the polypeptides;  
CC a storage medium suitable for use in an electronic device comprising  
CC information of the amino acid sequence of a polypeptide of the invention  
CC or the nucleotide sequence encoding the polypeptide; and a process  
CC comprising employing the polypeptide, or polynucleotide encoding it, in  
CC an industrial or household technical process. The present sequence is the  
CC protein sequence of a functional polypeptide of the invention.

Seq Sequence 248 AA;

Query Match % 100.0%; Score 1074; DB 9; Length 248;

Best Local Similarity 100.0%; Pred. No. 1.3e-89; Mismatches 207; Conservative 0; Indels 0; Gaps 0;

QY 1 NTVTMIRGTPRNASTVSAATTSIVNSTSSQVAKQONSSTSPARKSTNSLQHAQQAAT 60  
DB 42 NTVTMIRGTPRNASTVSAATTSIVNSTSSQVAKQONSSTSPARKSTNSLQHAQQAAT 101  
QY 61 TSSSSQKATVTPHYTKGKGLDLEIRVNSIQOVKSVGYDGIGETANCAFWVINTTINDGS 120  
DB 102 TSSSSQKATVTPHYTKGKGLDLEIRVNSIQOVKSVGYDGIGETANCAFWVINTTINDGS 161  
QY 121 TPEAEVVDGIFHLQNLNGVYQPDSTABIYANTNSGTITPDLNPGVSMTNLVFDMDFMT 180  
DB 162 TPEAEVVDGIFHLQNLNGVYQPDSTABIYANTNSGTITPDLNPGVSMTNLVFDMDFMT 221  
QY 181 YGHVGQHSIVLASMGPFGSDDETTALP 207  
DB 222 YGHVGQHSIVLASMGPFGSDDETTALP 248

RESULT 4  
AAU37553  
ID AU37553 standard; protein; 331 AA.

AC AAU37553;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #1723.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;  
KW antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.  
PR 23-OCT-2000; 2000US-0242578P.  
PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,  
PI Yamamoto RT, Xu HH;

DR WPI; 2001-611495/70.

XX N-PSDB; AAS55412.

PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.

PS Example 3; SEQ ID NO 13146; 511bp; English.

XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

Seq Sequence 331 AA;

Query Match % 12.6%; Score 135; DB 4; Length 331;

Best Local Similarity 31.0%; Pred. No. 0.0011; Mismatches 35; Conservative 25; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSIQOVKSVGYDGIGETANCAFWVINTTINDGSTPEAEVVDGIF 130  
DB 81 HIGETVKNKGLDLEIRVNSIQOVKSVGYDGIGETANCAFWVINTTINDGSTPEAEVVDGIF 140  
QY 131 HLQNLNGVYQPDSTABIYAN-TNSGTITPD-----LAPGVSMTNLVFDMDFMT 177  
DB 141 KIKS-GDKTFEADNTGMSANQSDNGSIENSFPLQRIINDSTAOGIQVDSV 192

RESULT 5

AAU34416  
ID AU34416 standard; protein; 331 AA.

AC AAU34416;

DT 14-FEB-2002 (first entry)

DE Staphylococcus aureus cellular proliferation protein #692.

XX Antisense; prokaryotic cellular proliferation protein; antibiotic;  
KW antibacterial; drug design.

OS Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

```

XX 21-MAR-2000; 2000US-019107AP.
PR 23-MAY-2000; 2000US-0206848P.
PR 26-MAY-2000; 2000US-0207727P.
PR 23-OCT-2000; 2000US-0242578P.
PR 27-NOV-2000; 2000US-0253625P.
PR 22-DEC-2000; 2000US-0257931P.
PR 16-FEB-2001; 2001US-0269308P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
PI Haeselbeck R, Ohlsen KU, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR MPI; 2001-611495/70.
DR N-PSDB; AAS52275.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids.
XX
XX Example 3; SEQ ID NO 5912; 511bp; English.
XX
PS The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 331 AA;
XX
XX Query Match 12.6%; Score 135; DB 4; Length 331;
XX Best Local Similarity 31.0%; Pred. No. 0.0011;
XX Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4
XX
QY 74 HTYG--KVGDEIRVNSLQOVKGYDGIETANGAFVINTIRNDGSTMEVVDGIF 130
DB 81 HRTGTVANGDLELVYVNSVETKMSQVPSLAPFNKAGIFVADVITINKKREKALTIDSSMF 140
QY 131 HLQNLGNVYQDSTAEIYAN-TNSGTIPTD-----LNDGVSMTNLVDFMDP 177
DB 141 KLKS-GDKTFEADNTGSMASQSDNGSIENSFFLQRLINDPSTAQGRIVPDSR 192
XX
XX RESULT 6
XX AAU37279
XX AAU37279 standard; protein; 331 AA.
XX
XX AAU37279;
XX
XX 14-FEB-2002 (first entry)
XX
XX Staphylococcus aureus cellular proliferation protein #1449.
XX
XX Antisense; prokaryotic cellular proliferation protein; antibiotic;
XX antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX

```

```

PD 27-SEP-2001.
XX
PR 21-MAR-2001; 2001M0-US009180.
XX
PR 21-MAR-2000; 2000US-0191078P.
XX
PR 23-MAY-2000; 2000US-0206848P.
XX
PR 26-MAY-2000; 2000US-0207727P.
XX
PR 23-OCT-2000; 2000US-0242578P.
XX
PR 27-NOV-2000; 2000US-0253625P.
XX
PR 22-DEC-2000; 2000US-0257931P.
XX
PR 16-FEB-2001; 2001US-0269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haeelbeck R, Ohlsen KJ, Zyckind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
XX
DR N-PsDB; AAS55138.
XX
PT New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids.
XX
PS Example 3; SEQ ID NO 12872; 51pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the genes,
CC their use in the discovery of novel antibiotics, the essential genes
CC themselves and the encoded proteins. The prokaryotes used are Escherichia
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC useful for the identification of potential new targets for antibiotic
CC development. The antisense nucleic acids can also be used to identify
CC proteins used in proliferation, to express these proteins, and to obtain
CC antibodies capable of binding to the expressed proteins. The proteins can
CC be used to screen compounds in rational drug discovery programmes. The
CC antisense nucleic acid sequence is also useful to screen for homologous
CC nucleic acids which are required for cell proliferation in a wide variety
CC of organisms. The present sequence represents an essential prokaryotic
CC cellular proliferation protein. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 331 AA;
XX
Query Match 12.6%; Score 135; DB 4; Length 331;
Beat Local Similarity 31.0%; Pred. No. 0.0011;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4
QY 74 HTYG---KVGDLEIRVNSLQGVKSYGVYDGIETANGAFWVINTIRNDGSPMEVVDI 130
DB 81 HKIGETVNGDLEIRVNSVETMKSGPSLAPTNKAGIVVADVTIKKNGKALTRIDSSWF 140
QY 131 HLQINNGVNYGPDSTAEIYAN-TNSGTPD-----LMPGVSMTNLVLPDMD 177
DB 141 KTKS-GDPTFEADNTGSMASQDSNGSIENSFFLQRIINPDSTAQGIYDVSE 192
XX
RESULT 7
XX
ID ABU16122
XX
AC ABU16122 standard; protein; 331 AA.
XX
AC ABU16122;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #1649.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Staphylococcus aureus.
XX

```





Qy	1	NTVVMIRKENTPRNST--VSAATSLVNSTSSQVAKQBNQNSTSPAHKSTNSLQHAQHA	58
Db	1323	NQNYIDADESRNATQAVTAABEGLNQTGANTSKADVNALVTYRAKALNGAENLR	1382
Qy	59	ATSSSSQSKLVIYFPHYTGKGDLEIRNLSLQQKSV-GYDGIETANGAFWVINITIRN	117
Db	1383	MTKXISATNTINGLPRNLTDQLQDNLKHQVBAQNVAGVGVDDKNTLANTMGAIRTSIQN	1442
Qy	118	DGSTPMEVVDGIFHLQNLNGNVYQPDSTAEIYVANTSGTIPDNLPGVSMTTNLVFDM	175
Db	1443	DNTT--KTSQVYLDASDSKNKNY---NTA---VNNANGVINVNTNP--NMDANANGM	1490
RESULT 11			
AAU37017	ID	AAU37017 standard; protein; 5795 AA.	
AAU37017;	AC		
14-FEB-2002	DT	(first entry)	
Staphylococcus aureus cellular proliferation protein #1187.	XX		
Antisense; prokaryotic cellular proliferation protein; antibiotic;	XX		
antibacterial; drug design.	KW		
Staphylococcus aureus.	XX		
MO200170955-A2;	PN		
27-SEP-2001.	PD		
21-MAR-2001; 2001WO-US0009180.	PF		
21-MAR-2000; 2000US-0191078P.	PR		
23-MAY-2000; 2000US-0206848P.	PR		
26-MAY-2000; 2000US-0207727P.	PR		
23-OCT-2000; 2000US-0242578P.	PR		
27-NOV-2000; 2000US-0253625P.	PR		
22-DEC-2000; 2000US-0257931P.	PR		
16-FEB-2001; 2001US-0269308P.	PR		
(BLIT-) BLITRA PHARM INC.	PA		
Haselbeck R, Ohlsen KT, Zyseknd JW, Wall D, Trawick JD, Carr GJ,	PI		
Yamamoto RT, Xu HH;	PI		
WPI; 2001-611495/70.	DR		
N-PSDB; AAS54876.	DR		
New polynucleotides for the identification and development of	PT		
antibiotics, comprise sequences of antisense nucleic acids.	PT		
Example 3; SEQ ID NO 12610; 511pp; English.	PS		
The invention relates to antisense inhibitors of genes essential to			
prokaryotic cellular proliferation, their use in identifying the genes,			
themselves and the encoded proteins. The prokaryotes used are <i>Escherichia</i>			
<i>coli</i> , <i>Staphylococcus aureus</i> , <i>Salmonella typhi</i> , <i>Klebsiella pneumoniae</i> ,			
<i>Pseudomonas aeruginosa</i> and <i>Enterococcus faecalis</i> . The invention is also			
useful for the identification of potential new targets for antibiotic			
development. The antisense nucleic acids can also be used to identify			
proteins used in proliferation, to express these proteins, and to obtain			
antibodies capable of binding to the expressed proteins. The proteins can			
be used to screen compounds in rational drug discovery programmes. The			
antisense nucleic acid sequence is also useful to screen for homologous			
nucleic acids which are required for cell proliferation in a wide variety			
of organisms. The present sequence represents an essential prokaryotic			
cellular proliferation protein. Note: The sequence data for this patent			
did not form part of the printed specification, but was obtained in			
electronic format directly from WIPO at			

CC	ftp.wipo.int/pub/published_poc_sequences
XX	
XX	SQ Sequence 5795 AA;
Query Match	10.0%; Score 107.5; DB 4; Length 5795;
Best Local Similarity	24.2%; Pred. No. 17;
Matches	43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;
OY	1 NTVMIRGNTPRNAST--VSAATSLVNSTNSQVAKOEONSSTSPAHKSTNLSQHACHOA 58
DB	3600 NQNTLADDESRRNAVTOAVTAABEGILNKQGCGNTSKADVNALNTVTARVAALNGAENLR 3659
OY	59 ATSSSQSKARIYPFHYYGKVGDLERIVNSLDQVKSY-GRDGGETANGAFWVINITRYN 117
DB	3660 NTKTSANTTJINGELPNMLTOLQDMIKHVEBAQNVAAGVNGVDKNGKITNTMTMGALETSION 3719
OY	118 DGSPTMEVVOGIFFHLONGNVQOPDSSTARIVANTSGTIPDLNPGVSMTTNVPFM 175
DB	3720 DNIT-KTSQNYLDASPSKNKY---NTA---VNNANGVINVTNPP--NPDAAALNGM 3767
RESULT 12	
ID	ADJ34822
AC	ADJ34822 standard; protein: 922 AA.
DT	22-APR-2004 (first entry)
DE	Xylanase from an environmental sample seq id 38.
KW	antibacterial; fungicide; thermostable xylanase activity;
KW	dough conditioning; beverage production; nutritional supplement;
KW	animal feed; lignin reduction; wood product; xylan; bacterial infection;
KW	fungal infection; coccidiosis.
OS	Unidentified.
PN	MO2003106654-A2.
PD	24-DEC-2003.
PF	16-JUN-2003; 2003MO-USO19153.
PR	14-JUN-2002; 2002US-0389299P.
PA	(DIVE-) DIVERSA CORP.
PI	Steer B, Callen W, Healey S, Hazlewood G, Wu D, Blum D,
PI	Eseghlalian A;
DR	WPI; 2004-099016/10.
N-PSDB:	ADJ34821.
Novel xylanase recombinant polypeptide useful for improving textile texture, treating paper, eliminating microorganisms.	
Claim 60; SEQ ID NO 38; 570pp; English.	
The invention describes an isolated or recombinant polypeptide (I), having 50% or more identity to 190 300-1200 residue amino acid sequences (S1), given in the specification, over a region of 100 or more residues and the polypeptide as thermostable xylanase activity. (I) is useful for: dough conditioning; beverage production; as a nutritional supplement in animal feed; reducing lignin in a wood or a wood product; and for eliminating and protecting animals from a microorganism comprising xylan. The polynucleotide (II) encoding (I) is useful for amplifying nucleic acid encoding a polypeptide having a xylanase activity which involves amplification of a template nucleic acid with a primer pair capable of amplifying (II) or its subsequence. (I) is useful for treating and preventing bacterial infection and fungal infection e.g. coccidiosis. This is the amino acid sequence of a xylanase protein isolated from an environmental sample.	





**This Page Blank (uspto)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:18:42 ; Search time 22.7473 Seconds  
(without alignments)

Title: US-10-784-592-43\_COPY\_42\_248  
 Defect name: 1074

Sequence: 1 NTVYMIRGNTPRNASTVSAT.....YSLVASMGGFFGSDETTYALP 207

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits' satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

### Listing first 45 summaries

Database : PIR\_80:★

```
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	173.5	16.2	215	2	H84032	hypothetical prote
2	135	12.6	331	2	H89965	conserved hypothet
3	120.5	11.2	542	2	S64030	probable membrane
4	116	10.8	215	2	S77663	multicopy phage x
5	113.5	10.6	6713	2	H89921	hypothetical protei
6	105	9.8	389	2	S53975	probable membrane
7	100.5	9.4	202	2	E97130	uncharacterized se
8	100	9.3	394	2	A37853	flagellin, 40k - F
9	99	9.2	2271	2	P90073	hypothetical prot
10	97	9.0	846	2	AD2612	conserved hypothet
11	97	9.0	848	2	E97454	hypothetical prote
12	94	8.8	1375	2	FL70345	dextranucrase (EC
13	93.5	8.7	498	1	FLBC	flagellin - Escheri
14	93.5	8.7	784	2	S26638	SPR-1 protein - hu
15	93	8.7	982	2	T13653	hypothetical prote
16	92.5	8.6	535	2	B84443	hypothetical prote
17	92.5	8.6	1228	2	A57384	multimerin, endotr
18	91.5	8.5	1449	2	T30552	glucosyltransfera
19	91	8.5	452	2	T40654	hypothetical prote
20	90.5	8.4	233	2	H86925	probable lipoprote
21	90.5	8.4	1365	2	T30832	limp1 protein - My
22	90	8.4	687	1	B69649	beta-galactosidase
23	90	8.4	1311	2	C84538	hypothetical prote
24	89.5	8.3	820	2	T17519	cell surface antidi
25	89.5	8.3	1449	2	T30857	glucosyltransfera
26	89	8.3	344	2	T40167	hypothetical prote
27	89	8.3	455	2	T48240	hypothetical prote
28	88.5	8.2	365	2	D70043	hypothetical prote
29	88.5	8.2	511	2	C86798	prophage p13 prote

30	88.5	8.2	633	2	C32053	parasporeal crystal
31	88.5	8.2	1770	2	S56221	hypothetical prote
32	88	8.2	1051	2	T18351	ImpI protein - Myc
33	87.5	8.1	385	2	T38113	hypothetical serin
34	87	8.1	345	2	B83737	pectate lyase BH06
35	87	8.1	817	2	T32442	hypothetical prote
36	86.5	8.1	410	2	B35670	protein-tyrosine k
37	86.5	8.1	493	2	G90604	hypothetical prote
38	86.5	8.1	544	2	B97166	flagellar basal bo
39	86.5	8.1	666	2	S50452	hypothetical prote
40	86	8.0	349	2	T41394	hypothetical serin
41	86	8.0	518	2	T09354	hypothetical prote
42	86	8.0	1121	2	S54504	hypothetical prote
43	85.5	8.0	532	2	B35631	spore germination
44	85.5	8.0	1237	2	D71850	probable outer mem
45	85.5	8.0	1277	2	S70306	hypothetical prote

## ALIGNMENTS

RESULT 1  
H84032  
hypothetical protein BH3064 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C:Accession: H84032  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saseki, R.; Masui, N.; Fuji, F.; Hiraiwa,  
Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: H84032  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-215 <SNO>  
A:Cross-references: UNIPROT:Q9KBE2; UNIPARC:UP1000000C40B3; GB:AF001517; GB:BA000004; NITD:  
A:Experimental source: strain C-125  
C:Genetics:  
C:Gene: BH3064

Query Match	Score	DB 2	Length
16.2%	173.5	2	215

Matches 55; Conservative 32; Mismatches 80; Indels 15; Gaps 8;

28 NSSQVAKQ-EQNSSTSPAHKSTNSLQHAQHQAAT--TSSSQSKRYIPHTYG---KVG 81

35 D\$SETA\$QEPPEAEV\$SADQ\$SENE\$SEEPPEAE\$GTEEDT\$EESAE\$EEDPIAGL\$GALKVGD 94

82 LEIRVNSLOQVKS VGYDIGETANGAFWVINITIRNDGSTPMEVBDGIFHLQNLGNV-Y 140

95 VFTANGISTAGSVG-DVLTAEAKGTFLLVDVTLKNEGSDSLVDSFFKLK--VGDVEX 151

141 WDS IAB IAN INSG I F IDNFGVSMI INLVDMF-DEM I GHVGUHSI SVASMGFFGS 199

**THE UNIVERSITY OF CHICAGO**

[illegible]

## RESULT 2

conserved hypothetical protein SA1618 [Imported] - *Staphylococcus aureus* (strain STAB1)

```
Accession: HA9965
;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
```

A.: Mizutani-III. Y.: Kobayashi. N.: Sawano. T.: Inoue. R.: Kaito. C.: Seki  
; Kuroda, M.; Uhta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui

Shida, T.; Hattori, M.; Ogabawara, N.; Hayashi, H.; Hiramatsu, K. *Cancer* 357: 1325-1349, 2001

Reference number: A89758; MUID:21311952; PMID:11418146  
 Title: Whole-genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Accession: H8965  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-331 <KUR>  
A:Cross-references: UNIPROT:Q99770, UNIPARC:UPI000000CADD4, GB:BA000018, PID:g13701593, F  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: SAL618

Query Match 12.6%; Score 135; DB 2; Length 331;  
Best Local Similarity 31.0%; Pred. No. 0.003;  
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 74 HTYG---KYGDLEIRVNSIQVKSVDGIGETANGAFVWVINTIRNDGSTPEVVDGIF 130  
DB 81 HKIGETVNGDLEIVNVNSVETMKSVGPLAPFNAGIFVAVDTTKKNGKALTTIDSSMF 140

QY 131 HLQNINGVYQPDSTAEIYAN-TNSGTIPTD-----LMPGVSMTTNLVFDMPD 177  
DB 141 KLKS-GDKTFEADNTGSMANQSDNGSIENSFPLQRIINDSTAQCKIVFVDSR 192

RESULT 3  
S64030  
Probable membrane protein YGH028c - Yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein G3661  
C:Species: Saccharomyces cerevisiae  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
A:Accession: S64030  
A:Residues: 1-542 <HEB>  
A:Cross-references: UNIPROT:P53189, UNIPARC:UPI00001356BA, EMBL:Z72550, NID:g1322499, P  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:SCW11; MIPS:YGL028C  
A:Cross-references: SGD:S0002996  
A:Map position: 7L  
A:Keywords: transmembrane protein  
F:8-24/Domain: transmembrane #status predicted <TMM>

Query Match 11.2%; Score 120.5; DB 2; Length 542;  
Best Local Similarity 25.4%; Pred. No. 0.076;  
Matches 51; Conservative 27; Mismatches 60; Indels 63; Gaps 8;

QY 9 NTPRNAATVSATTSIVNSTNSQVAKQKQNSSTSPAHKSTNSIQAHQAQATTS----- 62  
DB 217 NTPSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 273

QY 63 -----SSGSKLRYIPHTYGVKGDLEIRVNSIQVKSVDGIG-----ETANGAF 108  
DB 274 SAATSSPAIAYSPYNDGCKSADAVSSDLTIKSKISKIRVGTDCNSFFETQPA 333

QY 109 WVINTIRNDG---STPEVVDGIFHLQNLNGVYQPDSTAEI--YANTNSGTIPTDLN 162  
DB 334 VKLGIKI-NQGLYITSSGVDSD-----DSVTTLLQYQGTG----- 369

QY 163 PGVSMTNLVFDMPDFTYGH 183  
DB 370 -----KDVFPDITVGN 380

RESULT 4  
S77663  
multicopy phage resistance protein - Mycobacterium smegmatis  
C:Species: Mycobacterium smegmatis  
C:Date: 24-Oct-1998 #sequence\_revision 24-Oct-1998 #text\_change 09-Jul-2004  
C:Accession: S77663  
R:Baron, E. K.; Hatfull, G. F.  
Mol. Microbiol. 21, 159-170, 1996

A:Title: Characterization of a Mycobacterium smegmatis gene that confers resistance to p  
A:Reference number: S77663; MUID:97000357; PMID:8843442  
A:Accession: S77663  
A:Molecule type: DNA  
A:Residues: 1-215 <BAR>  
A:Cross-references: UNIPROT:Q50442, UNIPARC:UPI000009BEC0, EMBL:U50335, NID:g1477566, P  
C:Genetics:  
A:Gene: mpr

Query Match 10.8%; Score 116; DB 2; Length 215;  
Best Local Similarity 24.5%; Pred. No. 0.052;  
Matches 37; Conservative 33; Mismatches 59; Indels 22; Gaps 5;

QY 36 KQNSTPAHKSTNSIQAHQA-----TTSSGSKLRYIPHTYGVKGDLEIRVN 87  
DB 52 DDKQDTTAE--NTATYQAQAPAAAPTEGEVETATITPTPR-----DGKKEFVVR 99

QY 88 SLQQ-VKSVYDG-IGETANGAFVWVINTIRNDGSTPEVVDGIFHLQNLNGVYQPDST 145  
DB 100 SVEPELSVGNPFLNQAKQGFVITLSVQNIQDRPGGFSFSPNQKLPDTGSRSPETDTS 159

QY 146 AEIYANTNSGTIPTDLNPGVSMTNLVFDMP 176  
DB 160 AQIALDNNDIAVWNIINPGNTVDSLVDMP 190

RESULT 5  
B89921  
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
A:Accession: B89921  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizutani-Ut, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: B89921  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-6713 <KUR>  
A:Cross-references: UNIPROT:Q99U54, UNIPARC:UPI000011021A, GB:BA000018, PID:g13701232, P  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebha

Query Match 10.6%; Score 113.5; DB 2; Length 6713;  
Best Local Similarity 24.4%; Pred. No. 7.8;  
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;

QY 1 NTVTIRKNTPRNST--VSATTSIVNSTNSQVAKQKQNSSTSPAHKSTNSIQAHQAQ 58  
DB 899 NONVYDADESKRNATYQAVTAAGILKQGTGNSKAVDNALNAVTAKALNGAEVNR 958

QY 59 ATSSGSKLRYIPHTYGVKGDLEIRVNSIQVKSVDGIG-----GETANGAFVWINTI 115  
DB 959 NAKTSATNTINGLPVLQLOKONLKHQVEQANV--VGVNGVKDKGNTLNTAKMLKRSI 1016

QY 116 RNDGSTPEVVDGIFHLQNLNGVYQPDSTAEIYANTNSGTIPTDLNPGVSMTNLVFDM 175  
DB 1017 QNDNTT--KTSGNVYDAADSKKNV---NTA---VNNANGVINATNNP--NMDAALNDM 1066

RESULT 6  
S53975  
probable membrane protein YMR305c - Yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YMR9952.07c  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Oct-2004  
C:Accession: S53975  
R:Connor, R.; Churcher, C.M.  
submitted to the EMBL Data Library, Apr-11 1995

A:Reference number: S53969  
A:Accession: S53973  
A:Molecule type: DNA  
A:Residues: 1-389 <CON>  
A:Cross-references: UNIPROT:Q04951; UNIPARC:UPI0000135689; EMBL:Z49212; NID:g798940; PIR  
C:Genetics:  
A:Gene: SGD:SCW10; MIPS:YMR305C  
A:Cross-references: SGD:S0004921  
A:Map position: 13K  
C:Superfamily: family 17 glucosidase  
C:Keywords: transmembrane protein  
F:6-22/Domain: transmembrane #status predicted <TM>

Query Match 9.8%; Score 105; DB 2; Length 389;

Best Local Similarity 22.9%; Pred. No. 0.82;

Matches 53; Conservative 40; Mismatches 90; Indels 48; Gaps 8;

QY 5 MIRGMT-----PNAATVATTS-----LVNSTNSQVAKORNS----- 39  
DB 43 VTSNGSERIVPVENAVVAITTSSTVVAQAATTSTLEPTTSANVVTSSQQTSTLSSEAA 102  
QY 40 -----STSPAHSNSTSLQHOQAATSSSQSK-LRYIPFTYGVKVDLEIRVNSLQO 91  
DB 103 STVGSSTSSPSSSSSTSSSASSSISASGAKGITSPTVNDSCSKTAQVADLQ 162  
QY 92 VASVGVGIGETANGAFWYINITIRNDGSTPMEVVDGIFHLQNLNGNVYQPDSTAETVAN 151  
DB 163 L--TGFDPNRLYGVDCSQVENVLAQATSSQKFL--GIYYVDKIQDAVNTITKSAVSYGS 218  
QY 152 TNSGT---IPTDLNPGVSMVTNLVFPDMPFTYGVHGVHSLVASMGPFGS 199  
DB 219 WDDITTVSGNELVNGSGSATTTQVGR-----YVSTAKSLATSGAGTGS 261

## RESULT 7

E97130

uncharacterized secreted protein, homolog YXKC Bacillus subtilis [imported] - Clostridium

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 09-Jul-2004

C:Accession: E97130#

R:Rolling, J.; Breton, G.; Omlchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: E97130

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-202 <KUR>

A:Cross-references: UNIPROT:Q97HY8; UNIPARC:UPI00000CA32C; GB:AE001437; PIDN:AAK79832.1;

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC1868

Query Match 9.4%; Score 100.5; DB 2; Length 202;

Best Local Similarity 23.5%; Pred. No. 0.78;

Matches 40; Conservative 38; Mismatches 71; Indels 21; Gaps 6;

QY 44 AHKSTNSLQHOQAATSSSQSKLRYI-PFTYGVKVDLEIRVNSLQOVKSQ-GYVDIG 101  
DB 33 AKETTSATVQNNKATSTTDKDKKITVGBEGSGAMSKVLDTQBTNTIAGGGEN 92  
QY 102 ETANGAFWYINITIRNDGSTPMEVVDGIFHLQNLNGNV-YQPDSTA-----ETIYANT 152  
DB 93 KTTQGFVVVHLQMTKKNVNAQYBPDNFTLGDIKTKAQYKQDMKAGETANGAKTIYAD 152  
QY 153 NS-GITPTDLNPGVSMVTNLVFPDMPFTYGVHGVHSLVASMGPFGS 201  
DB 153 DSFPGVYDKVNPULSKQTYIVFEP-----TNFINANAVLHGGDD 193

## RESULT 8

A37853

flagellin, 40K - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 28-Jun-1991 #sequence\_revision 28-Jun-1991 #text\_change 26-Aug-1999

C:Accession: A37853

R:Tottem, P.A.; Lory, S.

J. Bacteriol. 172, 7188-7199, 1990

A:Title: Characterization of the type a flagellin gene from Pseudomonas aeruginosa PAK.

A:Reference number: A37853; MUID:91072275; PMID:2123866

A:Accession: A37853

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-394 <TOT>

A:Cross-references: UNIPROT:Q990Y4; UNIPARC:UPI000016FCA1; GB:M57501; NID:g151225; PIDN:AAA63458.1; PID

C:Superfamily: flagellin

Query Match 9.3%; Score 100; DB 2; Length 394;

Best Local Similarity 24.4%; Pred. No. 2.1;

Matches 51; Conservative 32; Mismatches 72; Indels 54; Gaps 10;

QY 9 NTPNASTVSA-----TTSLVNS-----TNSQVAKQONSST 41  
DB 13 NTQRMNNSASALNTSLQRLSTGSRINSKADDAAGIQIANRLTSQVNGINVAATKANDGI 72  
QY 42 SPA-----HKSTNSLQAOH---QAATSSSQSKLRYIPFTYGVKVDLEIRVNSLQO 91  
DB 73 SLAQTBEGALQOSTNQLQMRDLSLQANGSNSDERALN-----GEAQLOKELDR 125  
QY 92 VASVGVGIGETANGAFWYINITIRNDGSTPMEVVD-GIFHL-QNLNGNVYQPDSTAET 148  
DB 126 ISNTTTFGRKLDLDSFGVAFQV---GSAANEIISVIGIDMSABSLNKTYYKRDGGAV 182  
QY 149 VANTNSGITPTDLNPGVS--MTTNLVFDM 175  
DB 183 TAATASGTV--DIAIGITGSAVNVKVD 209

## RESULT 9

F90073

hypothetical protein SA2447 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004

C:Accession: F90073

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F90073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-2271 <KUR>

A:Cross-references: UNIPROT:Q990Y4; UNIPARC:UPI00000CABR3; GB:BA000018; PID:g13702612; I

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA2447

Query Match 9.2%; Score 99; DB 2; Length 2271;

Best Local Similarity 24.5%; Pred. No. 25;

Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVNIRGNTPRNASTVATSLVNST-----NSQVAKQONSSTSPAHS-----TNS 50  
DB 101 NTQSBTVGN--QNSTTIEASTADSTVTKSSSV--QTSNSDVTSSSKSKVTSTNS 156  
QY 51 LQHOQAATSSSQSKLRYIPFTYGVKVDLEIRVNSLQOVKSQYDYGIGETANGAFWY 110  
DB 157 TSNQGEKLTSTSESTSK-----NTTSSSDTKVA----- 186  
QY 111 INITIRNDGSTPMEVVDGIFHLQNLNGNVYQPDSTAETIYANTSGITPTDLNPGVSMVTN 170  
DB 187 -----STSTSTQPI-----NTSTNGSTNS--NNTSGSTTSSSVNLNTSTTS 226



QY 109 WTYNIT-----IRNDGSPMEVVDGIFHLQNLNGVYQPDSTAEIYANT---NSGTT 157  
Db 184 -YANIKRVNGKTYTYKEDGLQKN-----YALNINCKTFPFDETGALSNITLPSKKGN 236  
QY 158 PTDLPNGVSMTTNLVFDMPDFMTYGVGHQHY 188  
Db 237 TNNDNNTNSFAQYQYQYVS-TDVANFEHV-DHY 265

## RESULT 13

FLEC

Flagellin - Escherichia coli (strain K-12)

C:Species: Escherichia coli

A:Variety: strain K-12

C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C:Accession: A37249; J00018; I41270; H64955; A28187

R:KuwaJima, G.; Asaka, J.I.; Fujiwara, T.; Fujiwara, T.; Node, K.; Kondo, E.

J. Bacteriol. 168, 1479-1483, 1986

A:Title: Nucleotide sequence of the hag gene encoding flagellin of Escherichia coli.

A:Reference number: A37249; MUID:87057066; PMID:3536885

A:Accession: A37249

A:Molecule type: DNA

A:Residues: 1-498 &lt;KUM&gt;

A:Cross-references: UNIPARC:UP04949; UNIPARC:UP10000000AD; GB:M14358; NID:g146311; PIDN:

R:HanaFuja, T.; Sakai, A.; Tomioka, A.; Enomoto, M.

Mol. Gen. Genet. 216, 44-50, 1989

A:Title: Isolation and characterization of Escherichia coli hag operator mutants whose h

A:Reference number: J00018; MUID:89281489; PMID:2659972

A:Accession: J00018

A:Molecule type: DNA

A:Residues: 1-284, V', 286-498 &lt;HAN&gt;

A:Cross-references: UNIPARC:UP10000016F1P3; GB:X17440; NID:g41649; PIDN:CAA35488.1; PID:g

A:Experimental source: strain K-12

R:Szekely, E.; Simon, M.

J. Bacteriol. 155, 74-81, 1983

A:Title: DNA sequence adjacent to flagellar genes and evolution of flagellar- phase vari

A:Reference number: I41269; MUID:83238225; PMID:6505924

A:Accession: I41270

A:Molecule type: DNA

A:Status: translated from GB/EMBL/DBJ

A:Residues: 1-20 &lt;RES&gt;

A:Cross-references: UNIPARC:UP1000016F1P5; GB:J01607; NID:g146315; PIDN:AA92491.1; PID:

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shaq, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: H64955

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-498 &lt;BIAT&gt;

A:Cross-references: UNIPARC:UP10000000AD; GB:AE000285; GB:U00096; NID:g1788229; PIDN:AA

A:Experimental source: strain K-12, substrain MG1655

R:KuwaJima, G.

J. Bacteriol. 170, 3305-3309, 1988

A:Title: Construction of a minimum-size functional flagellin of Escherichia coli.

A:Reference number: A28187; MUID:88257060; PMID:3290204

A:Accession: A28187

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-15,140-402,489-498 &lt;KU2&gt;

A:Cross-references: UNIPARC:UP1000011F532; GB:M21445; NID:g146313; PIDN:AA23951.1; PID:

C:Comment: This is the principal protein component of the flagellar filament, which is a

C:Genetics:

A:Gene: fliC, hag

A:Map position: 42 min

C:Superfamily: flagellin

C:Keywords: flagellum

Query Match 8 7%; Score 93.5; DB 1; Length 498;  
Best Local Similarity 21.7%; Pred. No. 9.1;

Matches 59; Conservative 33; Mismatches 99; Indels 81; Gaps 11;  
QY 12 RNASTVSATSTSVNSTNSQVAKQONSTSPAHKSTNSLQHAQHQAATSSQSGLRYI 71  
Db 53 RFTSNIKGLTQAARNANDGISVAQTTEGALSEINNLCQRRELVQATGTGMSRDLSS1 112  
QY 72 P-----FHTYGVGDLEIRVNS-----LQGV--KSYGVYG 99  
Db 113 QDEIKSLRLEIDRVSGQCTPFGVNVAKNSKTIQVANDNOTTITDLKQIDAKTGLDG 172  
QY 100 IG-----ETANGAFVWINITI-----RNDGSPMEVVDGIFHLQNLNG 137  
Db 173 FSVKXNDVTTTSAPVTAFGATTNNIKLTGILTSTEARADTGTGNTASIEGVY---TDNG 229  
QY 138 NTVQP-----DSTAEIYANT--NSGTTP-----TDLPNGVSMT-----TNLVDP 174  
Db 230 NDYVAKITGQDNDGKYAVTAVNDGTVVWATGATANAATVTANTTATTTSGTFVQID 289  
QY 175 MPDPMTYGVGHQHYSLVASMGPFGSDETTYAL 206  
Db 290 NTAGSATNLTGA-VSLVYKLQDSKGNDDTYTAL 320

## RESULT 14

SPR-1 protein - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

C:Accession: S26638

R:Hagen, G.; Mueller, S.; Beato, M.; Sucke, G.

Nucleic Acids Res. 20, 5519-5525, 1992

A:Title: Cloning by recognition site screening of two novel GT box binding proteins: a t

A:Reference number: S26638; MUID:93087156; PMID:1454515

A:Accession: S26638

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-784 &lt;HAG&gt;

A:Cross-references: UNIPROT:Q02446; UNIPARC:UP10000135D7E; EMBL:X68561; NID:g38419; PIDN

C:Genetics:

A:Gene: GDB:SP4; SPR-1

A:Cross-references: GDB:136781

A:Map position: 2q31-2q31

C:Keywords: DNA binding; transcription regulation

Query Match 8 7%; Score 93.5; DB 2; Length 784;  
Best Local Similarity 34.0%; Pred. No. 17;

Matches 34; Conservative 17; Mismatches 32; Indels 17; Gaps 4;

QY 9 NTPRNASTV-----SATTSLVNSTNSQVAKQONSTSPAHKSTNSLQHAQH 56  
Db 306 NTTTSATSTMPESPSSSTCTTASTSLNSDPLVSSADTQGYASTS-ASSSRTTIESQT 364  
QY 57 QAAITSSQSGLRYIPFHTYGVGDLEIRVNSLQGVKSYG 96  
Db 365 PAATSESAQSSSQ--LHANG-MONQOQDSNLSLQGVQIVG 400

## RESULT 15

hypothetical protein 95B7.2 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: T13653

R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.

submitted to the EMBL data library, April 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17694

A:Accession: T13653

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-982 &lt;FER&gt;

A:Cross-references: UNIPARC:UP10000128461; EMBL:AL021728; NID:e1355643; PID:e1301389; PI



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 16:09:09 ; Search time 135.574 Seconds  
(without alignments)  
1077.233 Million cell updates/sec

Title: US-10-784-592-43\_COPY\_42\_248

Perfect score: 1074

Sequence: 1 NTVMIRGNTPRNASTVSAT.....YSLVSMGPFSGSDFTTALP 207

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Uniprot 05.80:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	173.5	16.2	215	2	09K8E2_BACHD
2	139	12.9	331	2	06G8D3_STPAS
3	139	12.9	331	2	08NVZ1_STAM
4	135	12.6	331	2	06GFQ7_STAPAC
5	135	12.6	331	2	05HEX9_STAPAC
6	135	12.6	331	2	07A4Z0_STAPAC
7	135	12.6	331	2	099T70_STAPAC
8	121.5	11.3	202	2	04MIR5_BACCE
9	120.5	11.2	542	1	SCW11_YEAST
10	119.5	11.1	170	2	08RB32_CHLTE
11	116	10.8	215	2	050442_MYSCM
12	113.5	10.6	6713	2	0931R6_STAPAC
13	113.5	10.6	6713	2	099U54_STAPAC
14	110	10.2	387	2	08G265_PSEAB
15	106.5	9.9	10746	2	06G6X3_STAPAC
16	105	9.8	389	1	SCW10_YEAST
17	105	9.8	2370	2	082R83_STAPAC
18	104.5	9.7	766	2	054PB6_DICDI
19	104	9.7	329	2	066YMI_PASMU
20	104	9.7	387	2	08G266_PSEAB
21	104	9.7	393	1	FLICA_PSEAB
22	104	9.7	394	2	053ZR9_PSEAB
23	104	9.7	394	2	053ZS1_PSEAB
24	103.5	9.6	923	2	07UKM5_PSEAB
25	103	9.6	713	2	06P857_ACTAD
26	102.5	9.5	1259	2	08G9X4_DICDI
27	102.5	9.5	9904	2	08NMQ6_STAPAC
28	100.5	9.4	202	2	097HY8_CIOAB
29	99.5	9.3	696	2	054Y58_DICDI
30	99	9.2	333	2	P72126_PSEAB
31	99	9.2	340	2	P72123_PSEAB

32	99	9.2	355	2	05Z159_NOCFA
33	99	9.2	1221	2	054LUB_DICDI
34	99	9.2	2261	2	05HC93_STAPAC
35	99	9.2	2271	2	07A362_STAPAC
36	99	9.2	2271	2	099QY4_STAPAC
37	99	9.2	2275	2	06G620_STAPAC
38	99	9.2	2275	2	08NUJ3_STAPAC
39	98.5	9.2	10498	2	05HPV8_STAPAC
40	98	9.1	1128	2	07NZT4_CHRVO
41	98	9.1	1145	2	09UR03_CANAL
42	98	9.1	1298	2	08G647_DICDI
43	97.5	9.1	1359	2	08R666_DICDI
44	97.5	9.1	1551	2	08IS13_DICDI
45	97.5	9.1	1557	2	054FP3_DICDI

ALIGNMENTS

RESULT 1					
ID	09K8E2_BACHD	PRELIMINARY;	PRT;	215 AA.	
AC	09K8E2;				
DT	01-OCT-2000 (Tremblrel. 15, Created)				
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)				
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)				
DE	BH3064 protein.				
GN	OrderedLocustNames=BH3064;				
OS	Bacillus halodurans.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_Taxid=86665;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=C-125 / JCM 9153;				
RX	MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;				
RA	Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,				
RA	Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,				
RA	Horikoshi K.;				
RT	"Complete genome sequence of the alkaliphilic bacterium Bacillus				
RT	halodurans and genomic sequence comparison with Bacillus subtilis."				
RL	Nucleic Acids Res. 28:4317-4331(2000).				
DR	EMBL; BA000004; BAB06783.1; -; Genomic_DNA.				
DR	PIR; H84032; H84032.				
KW	Complete proteome.				
SQ	SEQUENCE 215 AA; 23072 MW; 6A05DASDE7DC6358 CRC64;				
Query Match					
Best Local Similarity 30.2%; Score 173.5; DB 2; Length 215;					
Matches 55; Conservative 32; Mismatches 80; Indels 15; Gaps 8;					
QY	28 NSSQVAKQ-EQNSSTSPAHKSTNSIQHQAHOAAAT--TSSSQKLRVYIPHTYG---KVGD 81				
DB	35 DSSRFAQPEPPEABEVSADQSENESEBPEBAETEDTEBSAEEDPIAGIGBALKVGD 94				
QY	82 LEIRNYSQQQKSVYDIDIGETANCAFWVINITTINDSTPEWYVDGFLHNLNAGNV-Y 140				
DB	95 VFTFANGSTGASVG-DVLTAAKAKTFLIVDTTINBSDSITVDSPEFKK--GVDEY 151				
QY	141 QPDSTAIRIYAUNTSGTIFPTDLPNGVSMNTNLVFPDP--DFMTYGHGQHYSLVASMGPFS 199				
DB	152 DSDSAGXVYANBGADFITKANFGLELPKRVFVDPQVLDSDDI-----LINVQGFPGT 207				
QY	200 DB 201				
DB	208 QQ 209				
RESULT 2					
ID	06G8D3_STPAS	PRELIMINARY;	PRT;	331 AA.	
AC	06G8D3;				
DT	05-JUL-2004 (Tremblrel. 27, Created)				
DT	05-JUL-2004 (Tremblrel. 27, Last sequence update)				

```
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocustNames=SAS1720;
OS Staphylococcus aureus (strain MSSA476).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Knight M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,
RA Baeson N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571857; CAG43524.1; -, Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 331 AA; 35855 MW; 753AFB5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 12.9%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDGTGTANGAFWVITTRNDGSTPMFVVDGIF 130
DB 81 HKIGETVKNQDLEVTNVSVEYTKSVGPSIAPTNAGTFVADVITKNKKEKALTIDSSMF 140
QY 131 HLQNINGVYQPDSTAEIYANTN-SGTIPTD-----LNPVSMNTNLVFDMPD 177
DB 141 KTKS-GDITFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 3
Q8NVZ1 STAAW PRELIMINARY; PRT; 331 AA.
AC Q8NVZ1;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Hypothetical protein MW1738.
GN OrderedLocustNames=MW1738;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Yamagishi Y., Iwami N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL: BA000033; BAB95603.1; -, Genomic_DNA.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 331 AA; 35855 MW; 753AFB5FDAF50CF9 CRC64;

Query Match
Best Local Similarity 12.9%; Score 139; DB 2; Length 331;
Matches 36; Conservative 23; Mismatches 44; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDGTGTANGAFWVITTRNDGSTPMFVVDGIF 130
DB 81 HKIGETVKNQDLEVTNVSVEYTKSVGPSIAPTNAGTFVADVITKNKKEKALTIDSSMF 140
QY 131 HLQNINGVYQPDSTAEIYANTN-SGTIPTD-----LNPVSMNTNLVFDMPD 177
DB 141 KTKS-GDITFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192
```

```
DB 141 KTKS-GDITFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 4
Q6GF07 STAAW PRELIMINARY; PRT; 331 AA.
AC Q6GF07;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Putative membrane protein.
GN OrderedLocustNames=SAR1880;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=282459;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Knight M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barton A.,
RA Baeson N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagsels K.,
RA James K.D., Leonard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrall B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL: BX571856; CAG40870.1; -, Genomic_DNA.
KM Complete proteome.
SQ SEQUENCE 331 AA; 35850 MW; 82E9C61D2D59066 CRC64;

Query Match
Best Local Similarity 12.6%; Score 135; DB 2; Length 331;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

QY 74 HTYG---KVGDLIRVNSLQOVKSVGYDGTGTANGAFWVITTRNDGSTPMFVVDGIF 130
DB 81 HKIGETVKNQDLEVTNVSVEYTKSVGPSIAPTNAGTFVADVITKNKKEKALTIDSSMF 140
QY 131 HLQNINGVYQPDSTAEIYAN-TNSGTIPTD-----LNPVSMNTNLVFDMPD 177
DB 141 KTKS-GDITFEADNTGMSANQNDGSIENSFFLQRIINPDSTAQKIVFDVSE 192

RESULT 5
Q5HEX9 STAAW PRELIMINARY; PRT; 331 AA.
AC Q5HEX9;
DT 10-MAY-2005 (TREMBLrel. 30, Created)
DT 10-MAY-2005 (TREMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Conserved domain protein, putative.
GN OrderedLocustNames=SACOL1847;
OS Staphylococcus aureus (strain COL).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=93062;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15774886; DOI=10.1126/STB.187.7.2426-2438.2005;
RA Gill S.R., Fouts D.E., Archer G.L., Mongodin E.F., Deboy R.T.,
RA Ravel J., Paulsen I.T., Kolonay J.F., Brinkac L.M., Beaman M.J.,
RA Dodson R.J., Daugherty S.C., Madupu R., Angiuoli S.V., Durkin A.S.,
RA Haft D.H., Vamathevan J.J., Khouri H., Utterback T.R., Lee C.,
RA Dimitrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H.,
RA Hance J.R., Nelson K.E., Fraser C.M.;
RT "Insights on evolution of virulence and resistance from the complete
RT genome analysis of an early methicillin-resistant Staphylococcus
RT aureus strain and a biofilm-producing methicillin-resistant
RT Staphylococcus epidermidis strain.";
RL J. Bacteriol. 187:2426-2438(2005).
```







```
Db 72 INQPDAAVLFDVITVRNDKREARTIAP--FKLIDENGAEYESSNA--WSVDGSLGILD 127
QY 160 DLNPGVSMNTNLPFMDPMFTYGHVGOHSLVASMFPSSDE 201
Db 128 SLNPGYKRGYIVFDVFR-----GKHVXLEVGSGYSSDK 162

RESULT 11
050442 MYCSM
ID 050442_MYCSM PRELIMINARY; PRT; 215 AA.
AC 050442;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, last annotation update)
Mpr.
GN Name=mpf;
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxId=1772;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97000357; PubMed=843442;
RA Barom B.K., Hatfull G.F.;
RT "Characterization of Mycobacterium smegmatis gene that confers
RL resistance to phages I5 and D29 when overexpressed.";
DR Mol. Microbiol. 21:159-170(1996).
EMBL, U5035; AAB4151.1; -; Genom1C_DNA.
PIR, S77663; S77663.
SQ SEQUENCE 215 AA; 22960 MW; 7B4786D30BD6AB86 CRC64;

Query Match 10.8%; Score 116; DB 2; Length 215;
Best Local Similarity 24.5%; Pred. No. 0.27;
Matches 37; Conservative 33; Mismatches 59; Indels 22; Gaps 5;

QY 36 BONSSTPAHSTNSLQHQHOA-----TTSSSOSKLRYPHTYGVGDLEIRN 87
Db 52 DDKQTTTTRAE--NTRATYAQAPALAPYKVEPTTATINTPVR-----DKEKEFVR 99
QY 88 SLQO-VKSVGYDG-IGETANGAFWVYINITIRNDGSTPEMVDGIFHLQNLNGVYQPDST 145
Db 100 SVEPGLSEVGNPFLNQKQKQGFVITVLSVQNIIGRPFQSFPSNKLPTBRSRFTDTS 159
QY 146 AEIYANTNSGTTPTDLPNGVSMNTNLVFDMP 176
Db 160 AQIALDNDNDIAVMDININPNTVDVSLVYDMP 190

RESULT 12
093186 STAM
ID 093186_STAM PRELIMINARY; PRT; 6713 AA.
AC 093186;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, last annotation update)
Hypothetical protein ebnA.
GN Name=ebhA; OrderedLocusNames=SAV1434;
OS Staphylococcus aureus (strain M50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158878;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Ogasawara N., Hayashi K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi K., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
```

```
RL Lancet 357:1225-1240(2001).
DR EMBL, BA000017; BAB57596.1; -; Genomic_DNA.
DR SMR; Q93186; 1-125.
DR InterPro; IPR011439; DUF1542.
DR InterPro; IPR011490; FIVAR.
DR InterPro; IPR002988; GA.
DR InterPro; IPR006530; YD.
DR Pfam; PF07564; DUF1542; 8.
DR Pfam; PF07554; FIVAR; 44.
DR Pfam; PF07554; FIVAR; 44.
DR Pfam; PF01468; GA; 46.
DR TIGRfam; TIGR01643; YD repeat 2x; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 6713 AA; 722314 MW; BBC00536AC341BPS CRC64;

Query Match 10.6%; Score 113.5; DB 2; Length 6713;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;

QY 1 NTVVMIKNTPRNAST--VSATSLVNSTNSQVAKQEQNSSTSPAHKSTNSLQHQHOA 58
Db 899 NONVYDADSKRNAYTQAVTAEGILNQTGNTSKADVNDALNVTRAKAALNGAEKLR 958
QY 59 ATTSSSOSKLRYPHTYGVGDLEIRNLSQVKSVDGJ---GETANGAFWVINITI 115
Db 959 NAKTSATNTINGLPVLTLQKDNLKHQYBQANV--VGNYGVKDKNTLTMTMGALRTSI 1016
QY 116 RNDGSTPEMVDGIFHLQNLNGVYQPDSTAEIYANTNSGTTPTDLPNGVSMNTNLVFDMP 175
Db 1017 QNDNTT--KTSQNYLADSDSNQNY---NTA---VNANGVINATNRP--NMDAVALNDM 1066

RESULT 13
099054 STAM
ID 099054_STAM PRELIMINARY; PRT; 6713 AA.
AC 099054;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, last annotation update)
EbnA protein.
GN Name=ebhA; OrderedLocusNames=SA1267;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=158879;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi K., Hiramatsu K.,
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
```

```
Query Match 10.6%; Score 113.5; DB 2; Length 6713;
Best Local Similarity 24.4%; Pred. No. 38;
Matches 44; Conservative 37; Mismatches 82; Indels 17; Gaps 7;
```

```
QY 1 NTVMIRKNTFRANST--VSATTSLVNSTNSQVAKQONSTSPAKSTNSLQAOHQA 58
DB 899 NONTYDADBSKRNATTAQVTAABGLNKQGTGNTSKADVDAALNAVTAKALINAEHLR 958
QY 59 ATTSSSOKLARIYPHTYGVKQDLEIRVNSLQOVKSVYDGI--GETANGAFWVINITI 115
DB 959 NAKTSATYTINGLPHLTLOLQKONLKHQVEQAQNV--VGVNGVKDKGNTLTMTMGAALRTSI 1016
QY 116 RNDGSTPMEVVDGIFHLQNLNGVNPQPDSTAEIYANTNSGTPTDLPNGVSMTTNLVFD 175
DB 1017 QNDNTT--KTSQNYIADASDSKNKNY--NTA--VNNANGVINATNPN--NMDANAINDM 1066

RESULT 14
086265 PSEAE PRELIMINARY; PRT; 387 AA.
ID 086265;
AC 086265;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Flagellin.
GN Name=fljC.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97072180; PubMed=8914989; DOI=10.1016/0014-5793(96)01099-X;
RT Spangenberg C., Heuer T., Burger C., Tummeler B.;
RT "Genetic diversity of flagellins of Pseudomonas aeruginosa.";
RL FEBS Lett. 396:213-217(1996).
DR EMBL; L81146; AAC28556.1; -; Genomic_DNA.
DR HSRP; O67803; 10RY.
DR GO; GO:0009420; C:Flagellar filament (sensu Bacteria); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0001539; P:ciliary or flagellar motility; IEA.
DR InterPro; IPR001029; Flagellin_C.
DR InterPro; IPR001492; Flagellin_N.
DR Pfam; PF00700; Flagellin_C_1.
DR Pfam; PF00669; Flagellin_N_1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C_1.
KW Flagellum.
SQ SEQUENCE 387 AA; 39506 MW; 32E862D751F6142D CRC64;

Query Match 10.2%; Score 110; DB 2; Length 387;
Best Local Similarity 25.6%; Pred. No. 1.8;
Matches 53; Conservative 29; Mismatches 75; Indels 50; Gaps 9;

QY 9 NTPRNASVSA-----TTSLVNS-----TNSQVAKQONST 41
DB 13 NTOALNNSSALNTSLQRLSTGRINSKADDAAGIQTANRLTSCVNGNLATKNANDGI 72
QY 42 SPA-----HKSTSLQAOH---QAATTSSSOKLARIYPHTYGVKQDLEIRVNSLQO 91
DB 73 SLAQAGALQOSTNIIQMRPLSLQANGSNDSEFAL-----NGEYKQLO--KEIDR 125
QY 92 VKSVCYDGISETANGAFWVINITIINDGSTPMEVVD-GIFHL--ONLNGVNPQPDSTAEI 148
DB 126 ISNTTTFGGRKLDSFGVAFQV---GSAANEIISVIGIDEMASALNGTYFKADGGGAV 182
QY 149 YANTNSGTPTDLPNGVSMTTNLVFD 175
DB 183 TAATASGTVDIALGITGISTVAVKVM 209

RESULT 15
06GGX3 STAA PRELIMINARY; PRT; 10746 AA.
ID 06GGX3;
AC 06GGX3;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
```

```
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Very large surface anchored protein.
GN Name=ebh; Ordered locus names=SKR1447;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxId=282458;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;
RA Holden M.T.G., Fell E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,
RA Knight M.C., Foster T.J., Moore C.B., Hurst L., Atkin R., Barron A.,
RA Baeson N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Cotton C., Cronin A., Doggett J., Dowd L.,
RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jegerle K.,
RA James K.D., Jernard N., Line A., Mayes R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CA64044.1; -; Genomic_DNA.
DR SMR; 06GGX3; 3908-4032.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016042; P:lipid catabolism; IEA.
DR InterPro; IPR011439; DUF1542.
DR InterPro; IPR011490; FIVAR.
DR InterPro; IPR002988; GA.
DR InterPro; IPR005877; Gpos_YSIK.
DR InterPro; IPR006530; YD.
DR Pfam; PF07564; DUF1542; 8.
DR Pfam; PF07564; RIVAR; 59.
DR Pfam; PF01468; GA; 54.
DR Pfam; PF04650; YSIK_signal; 1.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 1.
DR TIGRFAMs; TIGR01168; YSIK_signal; 1.
KW Complete proteome.
SQ SEQUENCE 10746 AA; 1157156 MW; 8471C3E2EA3B0248 CRC64;

Query Match 9.9%; Score 106.5; DB 2; Length 10746;
Best Local Similarity 23.6%; Pred. No. 2.6e+02;
Matches 42; Conservative 35; Mismatches 88; Indels 13; Gaps 6;

QY 1 NTVMIRKNTFRANST--VSATTSLVNSTNSQVAKQONSTSPAKSTNSLQAOHQA 58
DB 4806 NONTYDADBSKRNATTAQVTAABGLNKQGTGNTSKADVDAALNAVTAKALINAGADNLR 4865
QY 59 ATTSSSOKLARIYPHTYGVKQDLEIRVNSLQOVKSV-GYDGISETANGAFWVINITI 117
DB 4866 NAKTSATYTINGLPHLTLOLQKONLKHQVEQAQNVAGVGDKNLTMTMGAALRTSI 4925
QY 118 DGSTPMEVVDGIFHLQNLNGVNPQPDSTAEIYANTNSGTPTDLPNGVSMTTNLVFD 175
DB 4926 DNTT--KTSQNYIADASDINKKNY--NTA--VNNANGVINATNPN--NMDANAINDM 4973
```

Search completed: April 7, 2006, 16:23:14  
Job time : 136.574 secs



; Sequence 19725, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19725  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19725

Query Match 9.7%; Score 104; DB 2; Length 417;  
Best Local Similarity 25.4%; Pred. No. 0.045; Indels 54; Gaps 11;  
Matches 53; Conservative 32; Mismatches 70;

QY 9 NTPRNASTVSA-----TTSLVNS-----TNSOVAKQONSGST 41  
DB 36 NTQRILNNSASALNTSLQRLSTGSRINSKXDAAGLQIANRLTSQVNGLVATKXNANGI 95  
QY 42 SPA-----HKSTSLQAHQ-----QAATSSQSKRLYIPHTYGVKGDLEIRNLSLQ 91  
DB 96 SLAQAGALQOSTYILQRMRLSLQSANSGNSDBERTAL-----NGEVKQLQ---KEIDR 148  
QY 92 VSVGVYDGGGTANGAFWYINITIRNDGSTPMGVVD-GIFHL--ONLNGVYQPDSTARI 148  
DB 149 ISNTTTFGRKLLDSFGYASFOV---GSANETISVIGIDKASASLNGTYFKADGGAGV 205  
QY 149 YANTNSGTIPDLPNGVS--MTTNLVFDM 175  
DB 206 TAATASGV--DIAIGTIGSAVNVKVD 232

RESULT 3  
US-10-172-502-18  
; Sequence 18, Application US/10172502  
; Patent No. 6841154  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
; FILE REFERENCE: P07263US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-172-502-18

Query Match 9.2%; Score 99; DB 2; Length 485;  
Best Local Similarity 24.5%; Pred. No. 0.18;  
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;  
QY 1 NTVVIRGNTPRNASTVATSLVNST-----NSSOVAKQONSGSTSPAHKS-----TNS 50  
DB 11 NTQSETVGN--QNSTTEASTADSTSVTKNSSSV--QTSNSDFTVSSSEKSTSTNS 66  
QY 51 LQAHQQAATSSSQSKRLYIPHTYGVKGDLEIRNLSLQGVKSVGYDGGTANGAFWV 110  
DB 67 TSNQEKLTSTSESTSSK-----NTTSSSDTKSYA----- 96

QY 111 INTIRNDGSTPMGVVDGIFHLQNLNGVYQPDSTARIYANTNSGTIPDLPNGVSMTN 170  
DB 97 -----STSTEQPL-----NTSTNGSTAS--NNTSQSTTSSSVNLKNTSTTS 136  
QY 171 LVFDMPEFTYGVKQHSYVASMCFPSGDETTVAL 206  
DB 137 TSTAPVKLTSTSRLL-----AMSTFASAAITTAIV 164

RESULT 4  
US-10-172-502-4  
; Sequence 4, Application US/10172502  
; Patent No. 6841154  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
; FILE REFERENCE: P07263US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2283  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-172-502-4

Query Match 9.2%; Score 99; DB 2; Length 2283;  
Best Local Similarity 24.5%; Pred. No. 1.9;  
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVIRGNTPRNASTVATSLVNST-----NSSOVAKQONSGSTSPAHKS-----TNS 50  
DB 113 NTQSETVGN--QNSTTEASTADSTSVTKNSSSV--QTSNSDFTVSSSEKSTSTNS 168  
QY 51 LQAHQQAATSSSQSKRLYIPHTYGVKGDLEIRNLSLQGVKSVGYDGGTANGAFWV 110  
DB 169 TSNQEKLTSTSESTSSK-----NTTSSSDTKSYA----- 198  
QY 111 INTIRNDGSTPMGVVDGIFHLQNLNGVYQPDSTARIYANTNSGTIPDLPNGVSMTN 170  
DB 199 -----STSTEQPL-----NTSTNGSTAS--NNTSQSTTSSSVNLKNTSTTS 238  
QY 171 LVFDMPEFTYGVKQHSYVASMCFPSGDETTVAL 206  
DB 239 TSTAPVKLTSTSRLL-----AMSTFASAAITTAIV 266

RESULT 5  
US-09-248-796A-15273  
; Sequence 15273, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 15273  
; LENGTH: 994  
; TYPE: PRT  
; ORGANISM: Candida albicans  
US-09-248-796A-15273

Query Match 9.0%; Score 97; DB 2; Length 994;

Best Local Similarity 22.3%; Pred. No. 0.85; Matches 55; Conservative 34; Mismatches 82; Indels 76; Gaps 9,

```

QY      1  NTVVMMRGMTPNRAVSVAATSLV---NSTNSSQVAKOEONSS-----TSPAHKST 48
Db      147  STINGQASLTNGTSLTTITTKPTSLVPPSTNSSRVSTMRNSSBBSMWLEKSSKLSS 206

QY      49  NSLOHQAHOAAITSSQ-----SKRYIPHTYGVKVDLEIRVNSLQOAVSVGYDI 100
Db      207  TSPINSSSIATSTSESELSAATTSDBSGLNSSSSSVELSSLSSEADSSSSSESVETGSS 266

QY      101  GETANGAFVIVINTINNDSTPHEVVDGIFHLQNLGNVYQPSPT---AEIVANTSGTI 157
Db      267  DETAS-----NYSGDLKALIDTNAPTVFAKRSR--I 295

QY      158  PTDLNGV-----SMTTNLVEDMDPFTYG--HVGQHSVLVAMSGF-----196
Db      296  PLTIAGVYDNNKGKIGIKNTKFTYTNLLGNQDFMWYPLPYGLYMSKTSYGYGAOVGHNNVSDR 355

QY      197  -FGSDET 202
Db      356  VFGSINT 362

```

RESULT 6  
 US-09-009-620-2  
 : Sequence 2, Application US/09009620A  
 : Patent No. 6127603  
 : GENERAL INFORMATION:  
 : APPLICANT: Nichols, Scott E.  
 : TITLE OF INVENTION: Substitutes for Modified Starches and  
 : FILE REFERENCE: 0357C  
 : CURRENT APPLICATION NUMBER: US/09/009,620A  
 : EARLIER FILING DATE: 1998-01-20  
 : EARLIER FILING DATE: 1995-06-07  
 : NUMBER OF SEQ ID NOS: 2  
 : SOFTWARE: PasteSeq for Windows Version 3.0  
 : SEQ ID NO 2  
 : LENGTH: 349  
 : TYPE: PRN  
 : ORGANISM: Streptococcus mutans  
 : US-09-009-620-2

Query Match	8.8%;	Score 94;	DB 2;	Length 349;
Best Local Similarity	21.3%;	Pred. No. 0.36;		
Matches 45;	Conservative 35;	Mismatches 87;	Indels 44;	Gaps 9;

Cy		10	TPRNASATSLVNSTNSQVAQEONSSS-----PAKSTNSLCHAGQAAT	60
Db		67	TATPDSATSATSOPTAIVTDNVSTTNSNTANTANFVVKPPTTTSBOAKTDSDKIIT	126
Cy		61	TSSSQSK----RYIPFH-----TYKVADLEIRVNSLOOVKSVDGYGIGETANGAF	108
Db		127	TSKAVNRLTLANGKFPFANNNTAHPKVTYDKIYPFKPIKGKLQPSLSODDIALLCN---	183
Cy		109	VWINT-----IRNDOSTEMEVDVGIFHLQNINGANYQPDSIAEYAHT---NSGTI	157
Db		184	-VKNIIRKYNKKYYKEEDGTLQKN-----YALNINGKTFPFDETGALSNNTLPSKKGINI	236
Cy		158	PTDLNPGVSMTNLVEPMDFPYTHGHNGHY	188
Db		237	TNNDTNSFAQYNQVIS--TDVANFEHV-DHY	265

RESULT 7  
US-09-210-361--4  
Sequence 4, Application US/09210361  
Patent No. 6284479  
GENERAL INFORMATION:  
APPLICANT: NICHOLS, Scott E.  
TITLE OF INVENTION: Substitutes for Modified Starches and

```

1  TITLE OF INVENTION: Latexes in Paper Manufacture
2  FILE REFERENCE: 03570R
3  CURRENT APPLICATION NUMBER: US/09/210,361
4  CURRENT FILING DATE: 1998-12-11
5  EARLIER APPLICATION NUMBER: 09/007,999
6  EARLIER FILING DATE: 1998-01-16
7  EARLIER APPLICATION NUMBER: 08/478,704
8  EARLIER FILING DATE: 1995-06-07
9  EARLIER APPLICATION NUMBER: 09/009,620
10 EARLIER FILING DATE: 1998-01-20
11 EARLIER APPLICATION NUMBER: 08/485,243
12 EARLIER FILING DATE: 1995-06-07
13 EARLIER APPLICATION NUMBER: 09/008,172
14 EARLIER FILING DATE: 1998-01-16
15 EARLIER APPLICATION NUMBER: 08/482,711
16 EARLIER FILING DATE: 1995-06-07
17 NUMBER OF SEQ. ID NOS.: 6
18 SOFTWARE: PasteSeq for Windows Version 3.0
19 SEQ ID NO 4
20 LENGTH: 1375
21 TYPE: PRN
22 ORGANISM: streptococcus mutans
23
24 US-09-210-361-4

```

Query Match	8.8%;	Score 94;	DB 2;	Length 1375;
Best Local Similarity	21.3%;	Pred. No. 2.8;		
Matches 45;	Conservative 35;	Mismatches 87;	Indels 44;	Gaps 91;

```

QY 10 TPRASATVATSTSLVNSSTNSQVAAQEOBNSNS-----PAKSTNSLQAQQAAT 60
Dy 67 TATDTSTRTASTSOPTATVTDNSSTTNSSTNTTANTANFVVKPPTTSBQAKTDSKXIIT 126
QY 61 TSSQSKL-----RYIPFH-----TYKVGDELRVNSLQOVSVGYDGIGETANGAF 108
Dy 127 TSKAVNRLTANGKVFVAPANNNTAHPRVTYDKIVPIKPIGKLGKLPSSLSDDDLALGN---- 183
QY 109 WVINIT-----IRNDGSTPRMEVVDGI FHLQNLNGNVYQPDSTAEIYAAT---NSGTI 157
Dy 184 -VKNIRKVGKRYYYKEDGTLOKN-----YALNNGKTFPFDEFGALSNNTLPSKKGIN 236
QY 158 PTDLNPQVSMVTNLVFDMPDEPMYTGAVGQHY 188
Dy 237 TNNNDNTNSPQAQYNQVYS--TDVANEHVA--DHY 265

```

```

1 RESULT 8
2 US-09-740-274-4
3 Sequence 4, Application US/09740274
4 Patent No. 6465203
5 GENERAL INFORMATION:
6 APPLICANT: Nichols, Scott B.
7 TITLE OF INVENTION: Glucan-containing Compositions and Paper
8 FILE REFERENCE: 0357CRD
9 CURRENT APPLICATION NUMBER: US/09/740,274
10 CURRENT FILING DATE: 2000-12-19
11 PRIOR APPLICATION NUMBER: 09/210,361
12 PRIOR FILING DATE: 1998-12-11
13 PRIOR APPLICATION NUMBER: 09/007,999
14 PRIOR FILING DATE: 1998-01-16
15 PRIOR APPLICATION NUMBER: 08/478,704
16 PRIOR FILING DATE: 1995-06-07
17 PRIOR APPLICATION NUMBER: 09/009,620
18 PRIOR FILING DATE: 1998-01-20
19 PRIOR APPLICATION NUMBER: 08/485,243
20 PRIOR FILING DATE: 1995-06-07
21 PRIOR APPLICATION NUMBER: 09/008,172
22 PRIOR FILING DATE: 1998-01-16
23 PRIOR APPLICATION NUMBER: 08/482,711
24 PRIOR FILING DATE: 1995-06-07
25 NUMBER OF SEQ ID NOS: 6
26 SOFTWARE: FastrSeq for Windows Version 3.0
27 SEQ ID NO 4
28 LENGTH: 1375

```

TYPE: PRT  
ORGANISM: streptococcus mutans  
US-09-740-274-4

Query Match  
Best Local Similarity 21.3%; Score 94; DB 2; Length 1375;  
Matches 45; Conservative 35; Mismatches 87; Indels 44; Gaps 9;

QY 10 TPRNASTVATSLVNSTNSQVAKOEONSTRS-----PAKSTNSLOHQAHOAT 60  
DB 67 TATDTSTATSATSQATATYTDVNSTNSTNTATANFVVEPTTSEAKTDNSDKITT 126  
QY 61 TSSSQSKL---RYIPFH-----TYKVGDLIRVNSLQOVKSVGYDGIETANGAF 108  
DB 127 TSKAVNRLTATGKFPVANNNTAHPKVTVDKIVIPKIKGLKQPSLSODDIALGN--- 183  
QY 109 WYINT-----IRNGSTPMEVVDGIFHLQNLNGNVQDPSTAEIYANT---NSGTI 157  
DB 184 -VKNIRVNGKYYKEDGTLOKN-----VALNNGKTFPFDETGALSNNTLPKKGNI 236  
QY 158 PTDLPNGVSMNTNLVFDMPDFWTYGHVGHY 188  
DB 237 TNDNTNSFAQYNOYYS-TDVANFEHV-DHY 265

RESULT 9  
US-09-270-767-31859  
Sequence 31859, Application US/09270767  
Patent No. 6703491

GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 31859  
LENGTH: 313  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-31859

Query Match  
Best Local Similarity 25.7%; Score 93.5; DB 2; Length 313;  
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 5 MIRGTPRNASTVATSL-----VNSTNSQVAKO---EONSTRSPAKSTNSLOH 53  
DB 142 VVSTVSTSLTNTASTASTIMGRKTDSTHSNSSGAGSQFYCELNSTSRHNSLDRDVH 201  
QY 54 AQHQA-ATTSSSQSKLRYIPFTYGVGDLIRVNSLQOVKSVGYDGIETANGAFWYN 112  
DB 202 HGHASLMSASSSVSNLSLSSNSGGR-----QSGKLTHSGMW-NGLKSHNGT----- 249  
QY 113 ITIRNDGSTPMEVVDGIFHLQNLNGNVQDPDS 144  
DB 250 -----TNGGSP-----HINAGLVQPS 266

RESULT 10  
US-09-270-767-47076  
Sequence 47076, Application US/09270767  
Patent No. 6703491

GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 47076

LENGTH: 313  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
US-09-270-767-47076

Query Match  
Best Local Similarity 25.7%; Score 93.5; DB 2; Length 313;  
Matches 39; Conservative 24; Mismatches 50; Indels 39; Gaps 7;

QY 5 MIRGTPRNASTVATSL-----VNSTNSQVAKO---EONSTRSPAKSTNSLOH 53  
DB 142 VVSTVSTSLTNTASTASTIMGRKTDSTHSNSSGAGSQFYCELNSTSRHNSLDRDVH 201  
QY 54 AQHQA-ATTSSSQSKLRYIPFTYGVGDLIRVNSLQOVKSVGYDGIETANGAFWYN 112  
DB 202 HGHASLMSASSSVSNLSLSSNSGGR-----QSGKLTHSGMW-NGLKSHNGT----- 249  
QY 113 ITIRNDGSTPMEVVDGIFHLQNLNGNVQDPDS 144  
DB 250 -----TNGGSP-----HINAGLVQPS 266

RESULT 11  
US-09-041-991A-8  
Sequence 8, Application US/09041991A  
Patent No. 6107278

GENERAL INFORMATION:  
APPLICANT: Schnepf, H. Ernest  
APPLICANT: Narva, Kenneth E.  
APPLICANT: Muller-Cohn, Judy  
TITLE OF INVENTION: Toxins Active Against Pests  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/041,991A  
FILING DATE: 13-MAR-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-709  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-041-991A-8

Query Match  
Best Local Similarity 23.0%; Score 93.5; DB 2; Length 633;  
Matches 42; Conservative 33; Mismatches 73; Indels 35; Gaps 8;

QY 1 NTVWIRKN-TPRNASTVATSLVNSTNSQVAKOEONSTRSPAKSTNSLOHQAHOQA 59  
DB 472 NNTAVHENGWTHIAPEDNTGFTTSPTHATVNNQOTFTFSKRGKNSDLSRFSQ---- 527  
QY 60 TTSSSQSKLRYIPFTYGVGDLIRVNSLQOVKSVGYDGIETANGAFWV---INITIR 116



Db 528 ----SNTARY- TLKNGNSVNLKRS-----SIGNSTIRTINGRYTASNVTTTN 576  
QY 117 NDGSTPMEVVDGIFHLQNLANGNVYOPDSTAEIYANTNSGTIPDLNPGVSMTTNLVDPMP 176  
Db 577 NDG-----VNDNGARFSDINIGNVASSNS--DVPDLIN--VTLNSGTQFDLM 620  
QY 177 DFM 179  
Db 621 NIM 623

RESULT 12  
US-09-608-533A-8  
Sequence 8, Application US/09608533A  
Patent No. 653464  
GENERAL INFORMATION:  
APPLICANT: Schuepf, H. Ernest  
Narva, Kenneth R.  
Miller-Cohn, Judy  
TITLE OF INVENTION: Toxins Active Against Pests  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: Florida  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/608,533A  
FILING DATE: 30-Jun-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/041,991  
FILING DATE: 13-MARCH-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: WA-709D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-608-533A-8

Query Match 8.7%; Score 93.5; DB 2; Length 633;  
Best Local Similarity 23.0%; Pred. No. 0.98;  
Matches 42; Conservative 33; Mismatches 73; Indels 35; Gaps 8;

QY 1 NTVYMRGN-TPRNASVTSATSLVNSTNSQVAKQONSSTSPAHKSTNSLQHQHQA 59  
Db 472 NNIVVHNGETMIHLAPEDNTGFTISPIHATQVNNQTRFPISEKFGNQSISRFGQ---- 527  
QY 60 TTSSSSSKLRYPFTHTYGVGLERVNSLQGVKSGYGVGIGETANGAFV---INITR 116  
Db 528 ----SNTARY- TLKNGNSVNLKRS-----SIGNSTIRTINGRYTASNVTTTN 576  
QY 117 NDGSTPMEVVDGIFHLQNLANGNVYOPDSTAEIYANTNSGTIPDLNPGVSMTTNLVDPMP 176  
Db 577 NDG-----VNDNGARFSDINIGNVASSNS--DVPDLIN--VTLNSGTQFDLM 620

QY 177 DFM 179  
Db 621 NIM 623

RESULT 13  
US-09-538-092-1254  
Sequence 1254, Application US/09538092  
Patent No. 6753314  
GENERAL INFORMATION:  
APPLICANT: Glot, Loic  
APPLICANT: Mansfield, Traci A.  
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
FILE REFERENCE: 15966-542  
CURRENT APPLICATION NUMBER: US/09/538,092  
CURRENT FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: 60/127,352  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/178,965  
PRIOR FILING DATE: 2000-02-01  
NUMBER OF SEQ ID NOS: 1387  
SOFTWARE: CuraPseqFormatter Version 0.9  
SEQ ID NO 1254  
LENGTH: 784  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (0) --(0)  
OTHER INFORMATION: Polypeptide Accession Number Q02446  
US-09-538-092-1254

Query Match 8.7%; Score 93.5; DB 2; Length 784;  
Best Local Similarity 34.0%; Pred. No. 1.4;  
Matches 34; Conservative 17; Mismatches 32; Indels 17; Gaps 4;

QY 9 NTPRNASV-----SATSLSVNSTNSQVAKQONSSTSPAHKSTNSLQHQHQA 56  
Db 306 NTVYMRGN-TPRNASVTSATSLVNSTNSQVAKQONSSTSPAHKSTNSLQHQHQA 59  
QY 57 QAAATSSSSSKLRYPFTHTYGVGLERVNSLQGVKSGY 96  
Db 365 PAATSSSSSKLRYPFTHTYGVGLERVNSLQGVKSGY 96

RESULT 14  
US-09-949-016-10170  
Sequence 10170, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10170  
LENGTH: 786  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-10170

Query Match 8.7%; Score 93.5; DB 2; Length 786;  
Best Local Similarity 33.0%; Pred. No. 1.4;

Matches 33; Conservative 17; Mismatches 33; Indels 17; Gaps 3;

QY 9 NTPRNASV-----SATSIVNSTSSQVAKOEONSSSTPAHKSTNSLQHAQH 56

DB 308 NTTTSASTMPSPPSSSTCTTASTSLTSSDTLVSSADTQGYASTS-ASSERTIESQT 366

QY 57 QAAATSSSQSLRYIPFTYGVGDLEIRVNSLQGVKSQV 96

DB 367 PATTESEAQSSSQLOP----NGMQNAQDQNSNLQGVQIVG 402

## RESULT 15

US-09-949-016-6805

; Sequence 6805, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6805

; LENGTH: 1228

; TYPE: PRT

; ORGANISM: Human

US-09-949-016-6805

Query Match 8.6%; Score 92.5; DB 2; Length 1228;

Best Local Similarity 22.0%; Pred. No. 3.3;

Matches 47; Conservative 27; Mismatches 79; Indels 61; Gaps 9;

QY 16 TVSATTSIVNST--NSSQVAKOEONSSSTPAHKSTN----- 49

DB 1018 TVNLTTLVLIGRTQRTMDNIYPEEYSSCS-RHPCONGGTCINGRTSFTCACRHPPTGDNC 1076

QY 50 SLOHAHQATSSSQSLRYIPF-----HTYGVGDLEIRVNSLQGVKSQVYDGIERT 103

DB 1077 TIKLYEENLAPDFSGSYRAPMVAFFASHTYGMTIPGPIIFNNL---DVNYGASTYP 1132

QY 104 ANGAF-----WVINITIRNDGS--TPMEVVDGI-----FHLQNLNGNVY----- 140

DB 1133 RTGKRIRIPYLVGYVRYKTTIESFAHISGFLVVDGIDKLAFAESENINSRHCDRVLTDAL 1192

QY 141 -OPDSTAEIYANTSGTIFPTDINPGVSMTTNLVF 173

DB 1193 LEINYQGEVWMLAKGTIPAKFPFVYTFESGYLLY 1226

Search completed: April 7, 2006, 16:25:20  
Job time : 31.4813 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd

OM protein - protein search, using SW model

Run on: **Apr 17, 2006, 16:40:31 ; Search time 100.998 Seconds**  
(without alignments)  
856.362 Million cell updates/sec

Title: US-10-784-592-43\_COPY\_42\_248

Perfect score:

Sequence: 1 NTVYMRGNTPRNA~~SV~~AT.....YSLVASM~~GF~~GSDE~~TY~~ALP 207

### Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1867569 beqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

```
Database : Published Applications AA_Main:
1: /cgn2_6/prodataa1/pubppaa/us07_PUBCOMB pep:*
2: /cgn2_6/prodataa1/pubppaa/us08_PUBCOMB pep:*
3: /cgn2_6/prodataa1/pubppaa/us09_PUBCOMB pep:*
4: /cgn2_6/prodataa1/pubppaa/us10A_PUBCOMB pep:*
5: /cgn2_6/prodataa1/pubppaa/us10B_PUBCOMB pep:*
6: /cgn2_6/prodataa1/pubppaa/us11_PUBCOMB pep:*
```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1074	100.0	248	5	US-10-754-592-43	Sequence 43, Apple
2	135	12.6	331	3	US-09-815-242-5912	Sequence 5912, Apple
3	135	12.6	331	3	US-09-815-242-12872	Sequence 12872, Apple
4	135	12.6	331	3	US-09-815-242-13146	Sequence 13146, Apple
5	135	12.6	331	3	US-10-282-1224-44046	Sequence 44046, Apple
6	113.5	10.6	6713	4	US-10-282-1224-43811	Sequence 43811, Apple
7	107.5	10.0	2086	3	US-09-815-242-5639	Sequence 5639, Apple
8	107.5	10.0	5795	3	US-09-815-242-12651	Sequence 12651, Apple
9	105.5	9.8	1020	4	US-10-437-963-131231	Sequence 131231, Apple
10	105	9.8	2386	4	US-10-156-761-7751	Sequence 7751, Apple
11	103	9.6	2344	3	US-09-815-242-12713	Sequence 12713, Apple
12	100	9.3	394	4	US-10-125-692-12	Sequence 12, Apple
13	100	9.3	394	5	US-10-991-347-14	Sequence 14, Apple
14	99	9.2	485	4	US-10-172-502-18	Sequence 18, Apple
15	99	9.2	485	6	US-11-020-509-18	Sequence 18, Apple
16	99	9.2	2261	5	US-10-470-0488-60	Sequence 60, Apple
17	99	9.2	2271	4	US-10-282-1224-43924	Sequence 43924, Apple
18	99	9.2	2283	4	US-10-172-502-4	Sequence 4, Apple
19	99	9.2	2283	6	US-11-070-509-4	Sequence 4, Apple
20	98.5	9.2	2398	4	US-10-282-1224-70176	Sequence 70176, Apple
21	98.5	9.2	6281	3	US-09-815-242-12996	Sequence 12996, Apple
22	98.5	9.2	10498	5	US-10-470-0488-44	Sequence 440, Apple
23	97.5	9.1	393	4	US-10-437-963-127450	Sequence 127450, Apple
24	97	9.0	194	4	US-10-282-1224-57709	Sequence 57709, Apple
25	95	8.8	982	6	US-11-037-143-4851	Sequence 4851, Apple
26	95	8.8	982	6	US-11-037-143-28995	Sequence 28995, Apple
27	94.5	8.8	632	4	US-10-040-906A-4	Sequence 4, Apple

## ALIGNMENTS

28	94.5	8.8	632	6	US-11-099-545-4	Sequence 4, Appl1
29	94.5	8.8	634	6	US-11-067-557-166	Sequence 166, Appl
30	94.5	8.8	634	6	US-11-067-557-224	Sequence 224, Appl
31	94	8.8	1375	5	US-09-740-274-4	Sequence 4, Appl
32	94	8.8	1375	4	US-10-383-930-35	Sequence 35, Appl
33	94	8.8	1375	5	US-10-797-892-35	Sequence 35, Appl
34	94	8.8	2712	5	US-10-282-122A-67070	Sequence 67070, A
35	93.5	8.7	603	4	US-10-425-114-42780	Sequence 42780, A
36	93.5	8.7	627	4	US-10-040-906A-6	Sequence 6, Appl1
37	93.5	8.7	627	6	US-11-098-545-6	Sequence 6, Appl1
38	93.5	8.7	632	6	US-11-067-557-2	Sequence 2, Appl1
39	93.5	8.7	633	6	US-11-067-557-32	Sequence 32, Appl
40	93.5	8.7	633	6	US-11-067-557-38	Sequence 38, Appl
41	93.5	8.7	661	4	US-10-451-467A-554	Sequence 554, Appl
42	93.5	8.7	926	4	US-10-106-605-143	Sequence 143, Appl
43	93.5	8.7	934	6	US-11-097-143-41301	Sequence 41301, A
44	93.5	8.7	1309	6	US-10-450-763-60725	Sequence 60725, A
45	92.5	8.6	571	4	US-10-760-701-45451	Sequence 45451, A

## RESULT 1

```

: Sequence 43, Application US/10784592
: Publication No. US20050147983A1
: GENERAL INFORMATION:
: APPLICANT: Wiltong, Reinhard
: APPLICANT: Oostergaard, Peter
: APPLICANT: Laessen, Soren
: TITLE OF INVENTION: POLYPEPTIDES OF ALICYCLOBACILLUS SP.
: FILE REFERENCE: 10406, 203-US
: CURRENT APPLICATION NUMBER: US/10/784,592
: CURRENT FILING DATE: 2004-02-23
: NUMBER OF SEQ ID NOS: 55
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 43
: LENGTH: 248
: TYPE: PRT
: ORGANISM: Alicyclobacillus sp.
: FEATURES:
: NAME/KEY: SIGNAL
: LOCATION: (1)..(41)
: FEATURES:
: NAME/KEY: mat peptide
: LOCATION: (42)..(248)
: OTHER INFORMATION: functional polypeptide
: US-10-784-592-43

```

### Query Match

100.0%; Score 1074; DB 5; Length 248;

Matches 207; Conserva

ctive 0; Mismatches 0; Indels 0; Gaps 0

QY	1	NTYMYRGNTPRNASTVSAATTSLVNSTNSQVAKOEKNSSTSPAKSSTNSLOHAQOAAAT	60
Db	42	NTYMYRGNTPRNASTVSAATTSLVNSTNSQVAKOEKNSSTSPAKSSTNSLOHAQOAAAT	10
QY	61	TSSSQSLRARIYPHTYGVKVDLEIRVNSLQOVKSVGYDGIGETGANGAFVINITIRNDGS	12
Db	102	TSSSQSLRARIYPHTYGVKVDLEIRVNSLQOVKSVGYDGIGETGANGAFVINITIRNDGS	16
QY	121	TPMEVVDGIFHLONLGNVYQDPDSTRELYAANTNSGTFPDLNPGVSMNTNLVFDMPDPMT	18
Db	162	TPMEVVDGIFHLONLGNVYQDPDSTRELYAANTNSGTFPDLNPGVSMNTNLVFDMPDPMT	22
QY	181	YGHVGOHYSLVASMGFGSDDETTYALP	207
Db	222	YGHVGOHYSLVASMGFGSDDETTYALP	248

```
/ Sequence 5912, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5912
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-5912
```

```
Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy      74 HTG---KVGDLIRVNSLQOVKSYGDIGETANGAFWVINITIRNDGSTEPMEVVDGIF 130
      81 HRIGETVKGDLLEVTVNSVETKWSGSLAPTNAGIFVADVITIKNGKCALTTIDSSMF 140
      131 HLQNLNGVYOPDSTAETIYAN-TNSGTIFPD-----LNPVSMTNLVFDMPD 177
      141 KTKS-GDKTFRADNTGMSANQSDNGSIENSFFLQRIINPDSTAQKIYVDVSE 192
Db
```

```
RESULT 3
US-09-815-242-12872
/ Sequence 12872, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
```

```
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12872
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-12872
```

```
Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;

Qy      74 HTG---KVGDLIRVNSLQOVKSYGDIGETANGAFWVINITIRNDGSTEPMEVVDGIF 130
      81 HRIGETVKGDLLEVTVNSVETKWSGSLAPTNAGIFVADVITIKNGKCALTTIDSSMF 140
      131 HLQNLNGVYOPDSTAETIYAN-TNSGTIFPD-----LNPVSMTNLVFDMPD 177
      141 KTKS-GDKTFRADNTGMSANQSDNGSIENSFFLQRIINPDSTAQKIYVDVSE 192
Db
```

```
RESULT 4
US-09-815-242-13146
/ Sequence 13146, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13146
/ LENGTH: 331
/ TYPE: PRT
/ ORGANISM: Staphylococcus aureus
/ US-09-815-242-13146
```

```
Query Match      12.6% Score 135; DB 3; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
```

```
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
Qy 74 HTYG--KVGDEIRVNSIQOVKSVGYDGIETANGAFVIVITIRNDGSTMEVVDGIF 130
Db 81 HKIGTIVKNGDLEIVNVSVEITKSVGPSLAPTNAKGIFVAVVITIKKCKEALTTIDSSWF 140
Qy 131 HLQNINGVYQPDSTAEIYAN-TNSGTIPTD-----LNPGVSMTNLVFDPMD 177
Db 141 KLS-GDKTFEADNTGMSANQSDNGSIENSFFLQRIINDSTAGQKIVDVSE 192

RESULT 5
US-10-282-122A-44046
; Sequence 44046, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44046
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-44046

Query Match 12.6%; Score 135; DB 4; Length 331;
Best Local Similarity 31.0%; Pred. No. 0.00055;
Matches 35; Conservative 25; Mismatches 43; Indels 10; Gaps 4;
Qy 74 HTYG--KVGDEIRVNSIQOVKSVGYDGIETANGAFVIVITIRNDGSTMEVVDGIF 130
Db 81 HKIGTIVKNGDLEIVNVSVEITKSVGPSLAPTNAKGIFVAVVITIKKCKEALTTIDSSWF 140
Qy 131 HLQNINGVYQPDSTAEIYAN-TNSGTIPTD-----LNPGVSMTNLVFDPMD 177
Db 141 KLS-GDKTFEADNTGMSANQSDNGSIENSFFLQRIINDSTAGQKIVDVSE 192
```

```
RESULT 6
US-10-282-122A-43811
; Sequence 43811, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43811
; LENGTH: 6713
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-43811

Query Match 10.6%; Score 113.5; DB 4; Length 6713;
Best Local Similarity 24.4%; Pred. No. 3.6; Indels 17; Gaps 7;
Matches 44; Conservative 37; Mismatches 82;
Qy 1 NTVYMRGNTFRNST--VSAITSLVNSTSSQVAKQONSSTSPAHKSTNSLQAHQQA 58
Db 899 NONVYLDADSKRNATVQAVTAABGILNKQGTGSKADVADNALNVTBAKALNCAENLR 958
Qy 59 ATTSSSQSKAYIRPHYTGXGDLIRVNSIQOVKSVGYDGI-----GERANGAFVIVITIR 115
Db 959 NAKTSATNTINGLPVLTQLOKDNLKHQVEQAQNV--VGVNGVKDKGNTLNTAMGALRTSI 1016
Qy 116 RNDGSTMEVVDGIFHLQNLQNGVYQPDSTAEIYANTNSGTIPTDLNPGVSMTNLVFDPMD 175
Db 1017 QNDMTT--KTSQNTYLDADSDSKNNY---NTA---VNNANGVYINATNNE--NMDANAINDM 1066

RESULT 7
US-09-815-242-5639
; Sequence 5639, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
```

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5639  
LENGTH: 2086  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-5639

Query Match 10.0%; Score 107.5; DB 3; Length 2086;  
Best Local Similarity 24.2%; Pred. No. 2.6;  
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

1 NTVMIRGNTPRNAST--VSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQQA 58  
1323 NONTVDADESKRNAYTQAVTAABGLNKQGTGNTSKADVNALNTVTAKAALNGAENLR 1382  
59 ATTSSQSKLRYPFHTYGVKGDLEIRVNSLQOVKSV-GYDGIETGANGAFVINITRN 117  
1383 NTKTSATNTINGLPVLQLOQDNLKHQVEQAQNVAGVGDKXGTLMTAKALNTSION 1442  
118 DQSTPMEVVDGIFHLQNLNGVNYQPDSTAEIYANTNSGTIPDLPVGSMTTNLVFDM 175  
1443 DNTT--KTSQNYLDASDSKNKNNY---NTA---VNNANGVINTVNNP--NMDANALNGM 1490  
Db

RESULT 8  
US-09-815-242-12610  
Sequence 12610, Application US/09815242  
Patent No. US2002061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12610  
LENGTH: 5795  
TYPE: PRT  
ORGANISM: Staphylococcus aureus  
US-09-815-242-12610

Query Match 10.0%; Score 107.5; DB 3; Length 5795;  
Best Local Similarity 24.2%; Pred. No. 11;  
Matches 43; Conservative 35; Mismatches 87; Indels 13; Gaps 6;

1 NTVMIRGNTPRNAST--VSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQQA 58  
3600 NONTVDADESKRNAYTQAVTAABGLNKQGTGNTSKADVNALNTVTAKAALNGAENLR 3659  
59 ATTSSQSKLRYPFHTYGVKGDLEIRVNSLQOVKSV-GYDGIETGANGAFVINITRN 117  
3660 NTKTSATNTINGLPVLQLOQDNLKHQVEQAQNVAGVGDKXGTLMTAKALNTSION 3719  
118 DQSTPMEVVDGIFHLQNLNGVNYQPDSTAEIYANTNSGTIPDLPVGSMTTNLVFDM 175  
3720 DNTT--KTSQNYLDASDSKNKNNY---NTA---VNNANGVINTVNNP--NMDANALNGM 3767  
Db

RESULT 9  
US-10-437-963-131231  
Sequence 131231, Application US/10437963  
Publication No. US20040123343A1

GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 131231  
LENGTH: 1020  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_33316C.1.dep  
US-10-437-963-131231

Query Match 9.8%; Score 105.5; DB 4; Length 1020;  
Best Local Similarity 25.1%; Pred. No. 1.5;  
Matches 51; Conservative 32; Mismatches 63; Indels 57; Gaps 11;

1 NTVMIRGNTPRNASTVSATSLVNSTSSQVAKQKQNSSTSPAHKSTNSLQHAQQA 60  
573 NTFYSSGK-----SASVLSMTLHSSQGS-----NSNMQIGQ----- 607  
61 TSSSQSKLRYPFHTYGVKGDLEIRVNSLQOVKSV-----VGYDGIETG--TANGAFW 109  
QY

Db 608 ----GGLKIQ-----YNAGALSTVSNSTKYVEANEQKKYKYYIGGDGKNSDPQASGAYI 658  
 QY 110 VI---NITRNDGSPFMEVYDGI---HLQNLGNVYQPDST--AEIYANTY--SGTIP 158  
 Db 659 FRPNGTVPFKTGQVFLIVLRGSIIDEVH--QQINPVIYQINRYKKDKDYETEFTVGPPI 717  
 QY 159 TDLPFGVSMNTNLVFDMPDPFTY 181  
 Db 718 VDDGNEKEKELSTEVYTNMAFNKTF 740

```

RESULT 10
US-10-156-761-7751
; Sequence 7751, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMTURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7751
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7751

```

	Query Match	9.8%	Score 105	DB 4	Length 2386
	Similarity	22.2%	Pred. No. 5.3		
	Matches	52	Conservative 30	Mismatches 92	Indels 60
				Gaps 10	
Qy	14	ASTVSATSLVNSTNSQVAKQEBNSSTSPHAKSTINSLOHAQHOAAATSSSQKLRYPF	73		
Db	1364	SKVVEHTSSAVYTTAYGYTAGGELAQIDPRGNNTLYYDMAAGRKTTDDPDAGLSSEY	1423		
Qy	74	HTYGGKGLERVNSLQOQKSVGVDIGIGETANG-----AFVYINITIRNDGSTPEWV	126		
Db	1424	NENQGVSTTATYTTADVQTVLTYYGYDNLRSATSVSGADELAAMWDDPAATGKG--QIT	1481		
Qy	127	DGIFHLQNLANGVNPQDSTAEIYANTNSG-----TIPTDLNPGVS--MTTNL	171		
Db	1482	SAV--SRASGNTY-----TTKKGKDKDERGRPLANTVILPTLVN-GLAGDYITTSV	1528		
Qy	172	VFDMPDFM-----TYGHVGQHSVLVANGPFGS---DETTY	204		
Db	1529	TYDADADHTTSVYPAGKLAAEKVTYYDDVGGQFRLRSSIG--GYVIYIDTTY	1580		

RESULT 11  
 US-09-815-242-12713  
 : Sequence 12713, Application US/09815242  
 : Patent No. US20020061569A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Haselbeck, Robert  
 : APPLICANT: Ohlsen, Karl L.  
 : APPLICANT: Zvekind, Judith W.  
 : APPLICANT: Wall, Daniel  
 : APPLICANT: Trawick, John D.  
 : APPLICANT: Yamamoto, Robert T.  
 : APPLICANT: Xu, H. Howard  
 : TITLE OF INVENTION: Identification of Essential Genes in

```

, TITLE OF INVENTION: Prokaryotes
, FILE REFERENCE: EILTRA.011A
, CURRENT APPLICATION NUMBER: US/09/815,242
, CURRENT FILING DATE: 2001-03-21
, PRIOR APPLICATION NUMBER: 60/191,078
, PRIOR FILING DATE: 2000-03-21
, PRIOR APPLICATION NUMBER: 60/206,848
, PRIOR FILING DATE: 2000-05-23
, PRIOR APPLICATION NUMBER: 60/207,727
, PRIOR FILING DATE: 2000-05-26
, PRIOR APPLICATION NUMBER: 60/242,578
, PRIOR FILING DATE: 2000-10-23
, PRIOR APPLICATION NUMBER: 60/253,625
, PRIOR FILING DATE: 2000-11-27
, PRIOR APPLICATION NUMBER: 60/257,931
, PRIOR FILING DATE: 2000-12-22
, PRIOR APPLICATION NUMBER: 60/265,308
, PRIOR FILING DATE: 2001-02-16
, NUMBER OF SEQ ID NOS: 14110
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 12713
, LENGTH: 2344
, TYPE: PRT
, ORGANISM: Staphylococcus aureus
US-09-815-242-12713

```

Query Match	9.6%	Score 103;	DB 3;	Length 2344;
Best Local Similarity	25.5%	Fred. No. 8;		
Matches	55;	Conservative 18;	Mismatches 71;	Indels 72;
			Gaps	9

  

Qy	1	NTVVMIRNTPRNASTVSAATSLVNST-----NSSQVAFQKONSSTSPAKS-----TNS	50
Db	101	NTQSEYVEN--QNSTTIDASTSTADSTSVTKNNSSV--QTSNSDPTVSSSKSNVNTSTNS	156
Qy	51	LQHAHQAAATSSSSQSKRIYIPFHTYKGVDELIRVNSLQOVKSYGYDGIGETANGAFW	110
Db	157	TSNOOEKLTSTSEBSRSK-----NTTSSSDPKSV-----	185
Qy	111	INTIRNDGSPFMEVVDGIPIHLQNLGNVYQDPSTAEIYANTNSGTFPTDLNPGVSMTN	170
Db	186	-----TSSSTSEQPI-----NTSNQSTNS--NNTSOSTTPIASNLKNTSTTS	226
Qy	171	LVFMDPDMFTYGVHGOHYSLVASKGFGSGDETTVAL	206
Db	227	TSTAPVKLTRESRL-----AMSTFSAATITLAL	254

```

RESULT 12
US-10-125-692-14
; Sequence 14, Application US/10125692
; Publication No. US20030044429A1
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan
; APPLICANT: Hayashi, Rumiaka
; APPLICANT: Smith, Kelly D.
; APPLICANT: Underhill, David M.
; APPLICANT: Ozinsky, Adrian
; TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods
; TITLE OF INVENTION: of Use
; FILE REFERENCE: P-US 5155
; CURRENT APPLICATION NUMBER: US/10/125,692
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: US 60/285,477
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 394
; TYPE: PRT
; ORGANISM: P. aeruginosa
US-10-125-692-14

```

Best Local Similarity 24.4%; Pred. No. 1.3;  
Matches 51; Conservative 32; Mismatches 72; Indels 54; Gaps 10;

QY 9 NTPRRASTVSA-----TSLVNS-----TNSQVAKOBONSST 41  
DB 13 NTKRNLNSSASLNTSLORLSTGRINSKADDAQLANRLTSGVNGLVATKNANGI 72  
QY 42 SPA-----HKSTNSLOHACH---QAAATSSSGSKLRYIPHTYGVGVDLEIRVNSIQ 91  
DB 73 SLAQTBAGALQOSTTILQRMRLSLQANGNSDSBERTALN-----GEAKQLKELDR 125  
QY 92 VKSVGYDGIGETANGAFVYVITITINDGSTPEVVD-GIFHL--ONLNGNVYQDPSTAEI 148  
DB 126 ISNTTFGGRKLLDGSFGVASFQV---GSAANEIISVIGIDEMASBSLNGTFKADGGGAV 182  
QY 149 YANTNSGTITPDLPNGVS--MTTNLVFDM 175  
DB 183 TPAATASGTV--DIALGITGSAVNKVKDM 209

RESULT 13  
US-10-991-347-14

; Sequence 14, Application US/10991347  
; Publication No. US20050147627A1  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan  
; APPLICANT: Hayaashi, Fumitaka  
; APPLICANT: Smith, Kelly D.  
; APPLICANT: Underhill, David M.  
; APPLICANT: Ozinsky, Adrian  
; TITLE OF INVENTION: Toll-Like Receptor 5 Ligands and Methods  
; FILE REFERENCE: 66661-122  
; CURRENT APPLICATION NUMBER: US/10/991,347  
; PRIOR FILING DATE: 2004-11-16  
; PRIOR APPLICATION NUMBER: US 60/285,477  
; PRIOR FILING DATE: 2001-04-20  
; PRIOR APPLICATION NUMBER: 10/125,692  
; PRIOR FILING DATE: 2002-04-17  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: P. aeruginosa  
US-10-991-347-14

Query Match 9.3%; Score 100; DB 5; Length 394;  
Best Local Similarity 24.4%; Pred. No. 1.3;  
Matches 51; Conservative 32; Mismatches 72; Indels 54; Gaps 10;

QY 9 NTPRRASTVSA-----TSLVNS-----TNSQVAKOBONSST 41  
DB 13 NTKRNLNSSASLNTSLORLSTGRINSKADDAQLANRLTSGVNGLVATKNANGI 72  
QY 42 SPA-----HKSTNSLOHACH---QAAATSSSGSKLRYIPHTYGVGVDLEIRVNSIQ 91  
DB 73 SLAQTBAGALQOSTTILQRMRLSLQANGNSDSBERTALN-----GEAKQLKELDR 125  
QY 92 VKSVGYDGIGETANGAFVYVITITINDGSTPEVVD-GIFHL--ONLNGNVYQDPSTAEI 148  
DB 126 ISNTTFGGRKLLDGSFGVASFQV---GSAANEIISVIGIDEMASBSLNGTFKADGGGAV 182  
QY 149 YANTNSGTITPDLPNGVS--MTTNLVFDM 175  
DB 183 TPAATASGTV--DIALGITGSAVNKVKDM 209

RESULT 14  
US-10-172-502-18  
; Sequence 18, Application US/10172502  
; Publication No. US20030185833A1  
; GENERAL INFORMATION:

; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .  
; FILE REFERENCE: P07263US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-172-502-18

Query Match 9.2%; Score 99; DB 4; Length 485;  
Best Local Similarity 24.5%; Pred. No. 2.1;  
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVMIRGNTPRNASTVSATSLVNST-----NSSQVAKOBONSSTSPARKS-----TNS 50  
DB 11 NTQSETVGN--QNSTTIRASTADSTSVTKNSSSV--QTSNSDVTSSSEKSKVTSTNS 66  
QY 51 LQHAQHQAATTSSSGSKLRYIPHTYGVGVDLEIRVNSLQOVKSVGYDGIGETANGAFV 110  
DB 67 TSNQGEKLTSTSESTSSK-----NTSSSDTKSYA----- 96  
QY 111 INTIRNDGSTPEVVDGIFHLQNLNGNVYQDPSTAEIYANTNSGTITPDLPNGVSMTTN 170  
DB 97 -----STSSTEQPI-----NTSTNQSTAS--NNTSOSTPSVNLKRTSTTS 136  
QY 171 LVFDMPEFTYGHVGHYSLVASMGEFGSDETTVAL 206  
DB 137 TSTAPVKLRTFSRL-----AMSTFASATTTAV 164

RESULT 15  
US-11-020-509-18

; Sequence 18, Application US/11020509  
; Publication No. US20050106648A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES. . .  
; FILE REFERENCE: P07263US02/BAS  
; CURRENT APPLICATION NUMBER: US/11/020,509  
; CURRENT FILING DATE: 2004-12-27  
; PRIOR APPLICATION NUMBER: US 10/172,502  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-11-020-509-18

Query Match 9.2%; Score 99; DB 6; Length 485;  
Best Local Similarity 24.5%; Pred. No. 2.1;  
Matches 53; Conservative 21; Mismatches 70; Indels 72; Gaps 9;

QY 1 NTVVMIRGNTPRNASTVSATSLVNST-----NSSQVAKOBONSSTSPARKS-----TNS 50  
DB 11 NTQSETVGN--QNSTTIRASTADSTSVTKNSSSV--QTSNSDVTSSSEKSKVTSTNS 66  
QY 51 LQHAQHQAATTSSSGSKLRYIPHTYGVGVDLEIRVNSLQOVKSVGYDGIGETANGAFV 110  
DB 67 TSNQGEKLTSTSESTSSK-----NTSSSDTKSYA----- 96  
QY 111 INTIRNDGSTPEVVDGIFHLQNLNGNVYQDPSTAEIYANTNSGTITPDLPNGVSMTTN 170  
DB 97 -----STSSTEQPI-----NTSTNQSTAS--NNTSOSTPSVNLKRTSTTS 136



OY 171 LVFDMDEWYGHVGOHYSLVASMGPFSGDETTYAL 206  
Db 137 TSTAPVKLRTFSRL-----AMSTFASAITTAV 164

Search completed: April 7, 2006, 16:44:39  
Job time : 100.998 secs

This Page Blank (uspto)





US-11-096-568A-22861  
 ; Sequence 22861, Application US/11096568A  
 ; Publication No. US20060048240A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexandrov, Nikolai et al.  
 ; TITLE OF INVENTION: Sequence-determined DNA Fragments and Corresponding Polypeptides  
 ; FILE OF INVENTION: Therby  
 ; FILE REFERENCE: 2750-1592PUS2  
 ; CURRENT APPLICATION NUMBER: US/11/096,568A  
 ; PRIOR FILING DATE: 2005-04-01  
 ; NUMBER OF SEQ ID NOS: 34471  
 ; SEQ ID NO 22861  
 ; LENGTH: 336  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays subsp. mays  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)..(336)  
 ; OTHER INFORMATION: Ceres Seq. ID no. 12410413  
 US-11-096-568A-22861

Query Match 8.4%; Score 90.5; DB 7; Length 336;  
 Best Local Similarity 19.0%; Pred. No. 0.61;  
 Matches 36; Conservative 30; Mismatches 72; Indels 51; Gaps 7;

QY 15 STVSATSLVNSTSSQVAKQKONSSTPAHAKSTNSLQAHQOAAATSSGSKRTYIPPH 74  
 DB 171 SVMPMDTAVRTNTNANVAVPSWMDTLAQPF-----SAGHVQGAALANNSSSM----- 218  
 QY 75 TVGKYVPLEIRVNSLQOVKSVGDIGETANGAFWVINITININDST-PMEVVDGIFHLQ 133  
 DB 219 -----ESPST-WFTSAVQENNVPELRAMPDPAQY 250  
 QY 134 NINGNVQDSTAEIYANTNSGTTPTDLNPGVSMTNLVFDM--PDFTYGVHGVGHYSIV 191  
 DB 251 NFIAGSIPDDTSGHL--QWLKAMDPIVDITALLIKRNLSMNLRSPPF-----EQRRRL 302  
 QY 192 ASMGPRGSD 200  
 DB 303 SSSYGGAD 311

## RESULT 6

US-11-130-821-4  
 ; Sequence 4, Application US/11130821  
 ; Publication No. US20060019275A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zuker, Charles S.  
 ; APPLICANT: Adler, Jon Elliot  
 ; APPLICANT: Lindemeier, Juergen  
 ; APPLICANT: Cowan, David  
 ; TITLE OF INVENTION: The Regents of the University of California  
 ; TITLE OF INVENTION: Nucleic Acid Encoding Proteins Involved in Sensory  
 ; FILE REFERENCE: 02307E-084210US  
 ; CURRENT APPLICATION NUMBER: US/11/130,821  
 ; CURRENT FILING DATE: 2005-05-16  
 ; PRIOR APPLICATION NUMBER: US/09/361,630  
 ; PRIOR FILING DATE: 1999-07-27  
 ; PRIOR APPLICATION NUMBER: US 60/094,464  
 ; PRIOR FILING DATE: 1998-07-28  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 4  
 ; LENGTH: 729  
 ; TYPE: PRT  
 ; ORGANISM: Mus sp.  
 ; FEATURE:  
 ; OTHER INFORMATION: mouse taste cell polypeptide (TCP) #2 amino acid  
 US-11-130-821-4

Query Match 8.3%; Score 89.5; DB 7; Length 729;

Best Local Similarity 28.5%; Pred. No. 2.2;  
 Matches 37; Conservative 14; Mismatches 44; Indels 35; Gaps 6;

QY 1 NTVMIRGNTPRNAS-----TVSATSLVNSTN-SSQVAKQKONS-----TSPAHKSTNS 50  
 DB 243 NTVALDTPGVSRGADSPQTPTSTTDSFKTSNLPQALQPSHGMLFTSPIHPTLS 302  
 QY 51 LQH-----AQHQAATTSSS-----QSKRTYIPHTYK---VGDLEIR 85  
 DB 303 LQHFSSPSTASSSGFTSSVHADPTLASLYLPHQQDMSLQDLSPTSGRSHHTHSVTR 362  
 QY 86 VNSLQOVSV 95  
 DB 363 INSNRPTKAV 372

## RESULT 7

US-11-052-554A-216  
 ; Sequence 216, Application US/11052554A  
 ; Publication No. US20050288866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
 ; FILE REFERENCE: 30853/40359A  
 ; CURRENT APPLICATION NUMBER: US/11/052,554A  
 ; CURRENT FILING DATE: 2005-02-07  
 ; PRIOR APPLICATION NUMBER: US 60/589,227  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
 ; PRIOR FILING DATE: 2004-02-06  
 ; NUMBER OF SEQ ID NOS: 763  
 ; SOFTWARE: Patent In version 3.3  
 ; SEQ ID NO 216  
 ; LENGTH: 507  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus mutans UA159  
 US-11-052-554A-216

Query Match 8.1%; Score 87.5; DB 7; Length 507;  
 Best Local Similarity 21.1%; Pred. No. 2.1;  
 Matches 56; Conservative 38; Mismatches 93; Indels 79; Gaps 14;

QY 2 TVTMIRGNTPRNASTVSATTS--LVNSTSSQVAKQKONSSTPAHKT-----NSLQH 53  
 DB 48 TVSQADGNPQGTTSVQOETAPOCTKTSQSSDATVDSBSATSPDEQTVSGNDNSSSQ 107  
 QY 54 AQHQAATTSSSQSKRTYIPHTYKGVGDLEIRVNSLQ----- 90  
 DB 108 IDQTIADNRDSD--HISKTSAAITTEDQEKVNSAKQTAATNNOPTRYSAKDAYGNS 165  
 QY 91 -----QVKSQYDQIGETANGAFWVINITININDOSTPM-----EVV-----DG 128  
 DB 166 NFNKTLTFGRKANVADVTVNGVD-----EIVVNDPSAPYVRNAEIAKLEY 216  
 QY 129 IFHLQNLGNVYQF-----DSTAEIYANTNSGTTPTDLNPGVSMTNLVFDMDFMTY--- 181  
 DB 217 LTELKLNIN-NIAIPVPSVDQWQKXAQDRAHEANEKQ-GLDHDTNL--PIPNNLTWVAE 272  
 QY 182 -GHVGQHSVLAASM--GF-FGSDETT 203  
 DB 273 DGHLDMSISQSKSQEGYTLASDKAT 298

## RESULT 8

US-11-052-554A-95  
 ; Sequence 95, Application US/11052554A  
 ; Publication No. US20050288866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sachdeva, et al.  
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE  
 ; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL  
 ; FILE REFERENCE: 30853/40359A

; CURRENT APPLICATION NUMBER: US/11/052,554A  
 ; CURRENT FILING DATE: 2005-02-07  
 ; PRIOR APPLICATION NUMBER: US 60/587,227  
 ; PRIOR FILING DATE: 2004-07-20  
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004  
 ; PRIOR FILING DATE: 2004-02-06  
 ; NUMBER OF SEQ ID NOS: 763  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 95  
 ; LENGTH: 1237  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori J99  
 ; US-11-052-554A-95

Query Match 8.0%; Score 85.5; DB 7; Length 1237;  
 Best Local Similarity 21.4%; Pred. No. 11;  
 Matches 48; Conservative 33; Mismatches 58; Indels 85; Gaps 11;

QY 17 VSATSLVNSTSSQVAKOE-----ONSSTSPAKSTNSLO---HAHQ-----A 58  
 DB 482 VSQVSVYNSLKTSENIQNAKILCNNGSGSTSPCNSSGGLISGNAQLQNTLSPTN 541  
 QY 59 ATTSSQ-----SKRYIPFHYGVGLDIRVNSLQOVKSVGYD-----GIGETAN 105  
 DB 542 GTTNTQAKSNASKLKRAM-----VWVNEBEAKTTFNQGSGPTTQSSNSTVM 589  
 QY 106 GAFWVINITIRNDGSTPEV-----VDGIFHLQNLNGVYQPDST----- 145  
 DB 590 GA-----LNTVLOVNSFPQSGISGAFQNGENNTQAMNALTNSNPNGOSQNLITNNODL 646  
 QY 146 -----AEIY--ANTNSGTFPTDLN-----PGVSMITTN 170  
 DB 647 RIQLAFNFYQLINTINQVPTMNLINQSQOTQOTSGSASTTN 690

## RESULT 9

US-11-087-099-3488  
 ; Sequence 3488, Application US/11087099  
 ; Publication No. US20060041961A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abad, Mark S. et al.  
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement  
 ; FILE REFERENCE: 38-21(53450)B EP  
 ; CURRENT APPLICATION NUMBER: US/11/087,099  
 ; CURRENT FILING DATE: 2005-03-22  
 ; NUMBER OF SEQ ID NOS: 12464  
 ; SEQ ID NO 3488  
 ; LENGTH: 500  
 ; TYPE: PRT  
 ; ORGANISM: Clostridium acetobutylicum  
 ; US-11-087-099-3488

Query Match 7.9%; Score 85; DB 7; Length 500;  
 Best Local Similarity 25.1%; Pred. No. 3.5;  
 Matches 42; Conservative 22; Mismatches 65; Indels 38; Gaps 5;

QY 10 TPRNST-----VSATSLVNSTSSQVAKOEONSSTSPAKSTNSLOHAQQAATSS 64  
 DB 199 TFSNASKNKNMTSSDINSIVSDNSGKITATAGSGTATA-----TTTAASS 245  
 QY 65 QSKLRYIPFHYGVGLDIRVNSLQOVKSVGYDGETANGAFWVINITIRNDGSTPE 124  
 DB 246 DSQKTATCCTVTVNKKRTPITFKDNNL-----EIS-----IRMNINKPTGSLVKE 288  
 QY 125 VVDGIFHLQNLNGVYQPD-----STAEIYANTNSGTFPTDLNPGVSMIT 168  
 DB 289 DVQGITSTFYNTNPNIHLYDGIENLSNLEFLTFSNTPIDLTPLKSLT 335

RESULT 10  
 US-10-055-877-140  
 ; Sequence 140, Application US/10055877  
 ; Publication No. US20050288241A1

## GENERAL INFORMATION:

; APPLICANT: Decristofaro, Marc  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Tcherniev, Vellizar  
 ; APPLICANT: Zhong, Mei  
 ; APPLICANT: Anderson, David  
 ; APPLICANT: Ballinger, Robert  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Spytek, Kimberly  
 ; APPLICANT: Kaceli, Luca  
 ; APPLICANT: Kerkuda, Rameesh  
 ; APPLICANT: Guo, Xiaojia  
 ; APPLICANT: Zernhusen, Bryan  
 ; APPLICANT: Andrew, David  
 ; APPLICANT: Mezas, Peter  
 ; APPLICANT: Paturajan, Meera  
 ; APPLICANT: Burgess, Catherine  
 ; APPLICANT: Eileen, Andrew  
 ; APPLICANT: Wolenc, Adam  
 ; APPLICANT: Baumgartner, Jason  
 ; APPLICANT: Shinkels, Richard  
 ; APPLICANT: Gusev, Vladimir  
 ; APPLICANT: Vermet, Corine  
 ; APPLICANT: Taupier Jr., Raymond  
 ; APPLICANT: Pena, Carol  
 ; APPLICANT: Shenoy, Suresh  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Caeman, Stacie  
 ; APPLICANT: Boldog, Ference  
 ; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
 ; FILE REFERENCE: 21402-251  
 ; CURRENT APPLICATION NUMBER: US/10/055,877  
 ; CURRENT FILING DATE: 2002-01-22  
 ; PRIOR APPLICATION NUMBER: 60/262,892  
 ; PRIOR FILING DATE: 2001-01-19  
 ; PRIOR APPLICATION NUMBER: 60/263,598  
 ; PRIOR FILING DATE: 2001-01-23  
 ; PRIOR APPLICATION NUMBER: 60/263,799  
 ; PRIOR FILING DATE: 2001-01-24  
 ; PRIOR APPLICATION NUMBER: 60/264,117  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/264,139  
 ; PRIOR FILING DATE: 2001-01-25  
 ; PRIOR APPLICATION NUMBER: 60/264,478  
 ; PRIOR FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: 60/263,351  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: 60/272,870  
 ; PRIOR FILING DATE: 2001-03-02  
 ; PRIOR APPLICATION NUMBER: 60/275,990  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/275,927  
 ; PRIOR FILING DATE: 2001-03-14  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 512  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 140  
 ; LENGTH: 1765  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-055-877-140

Query Match 7.9%; Score 85; DB 6; Length 1765;  
 Best Local Similarity 24.6%; Pred. No. 20;  
 Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;

QY 18 SATSLVNSTSSQVAKOEONSSTSPAKSTNSLOHAQQAATSSQSKLRYIPFHYG 77  
 DB 670 NANNVLSKSGSLTFLHLAAQDRVNAEVLVNGCAHV-----DAOTKGTTPHLVGC 721  
 QY 78 KVGDLEIRVNSLQOVKSV-----GYDGIGETA-NGAFWVINITIRNDGSTPEVVDGI 129

Db 722 HYGNIKI-VNFILOHSAKVNAKTKNGYTAHQAOQGHTHIINVLLQNNAS-PNEL---- 775  
QY 130 FHLQNLNG-----VYQPS-----TAIYANTNSGTPTDLPNGVSWTTLVDP 174  
Db 776 ----TVNGNTALAIARRLGIVSVDTLKVYTBEMITT---TTEKHKRNVPETNMEVLD 828  
QY 175 MPD 177  
Db 829 MSD 831

RESULT 11  
US-10-055-877-141  
; Sequence 141, Application US/10055877  
; Publication No. US20050288241A1  
; GENERAL INFORMATION:  
; APPLICANT: Decristofaro, Marc  
; APPLICANT: Padigaru, Murajidhara  
; APPLICANT: Miller, Charles  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Zhong, Mei  
; APPLICANT: Anderson, David  
; APPLICANT: Ballinger, Robert  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Ratelli, Luca  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Zernhusen, Bryan  
; APPLICANT: Andrew, David  
; APPLICANT: Mezes, Peter  
; APPLICANT: Paturajan, Meera  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Eissen, Andrew  
; APPLICANT: Wolenc, Adam  
; APPLICANT: Baumgartner, Jason  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Verneet, Corine  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Pena, Carol  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ference  
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby  
; FILE REFERENCE: 21402-251  
; CURRENT APPLICATION NUMBER: US/10/055,877  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,598  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: 60/263,799  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 60/264,117  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,139  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/264,478  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/263,351  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: 60/272,870  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/275,990  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/275,927  
; PRIOR FILING DATE: 2001-03-14  
; Remaining prior Application data removed - See file wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 512  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 141

; LENGTH: 1940  
; TYPE: PRF  
; ORGANISM: Mus musculus  
US-10-055-877-141

Query Match  
Best local Similarity 24.6%; Pred. No. 23;  
Matches 45; Conservative 24; Mismatches 70; Indels 44; Gaps 9;

QY 18 SATSLVNSTSSQVAKQEQNSSTSPAKSTNSLQHAQOAMTSSQSKRYIPHTYG 77  
Db 670 NANVNLNKSGLTPLHLAAQEDRVNVAEVLVNGAHV-----DAQTKMGYTPHLVGC 721  
QY 78 KYGDLKIRVNSLQYKVS-----GYDQIGETA-NGATWVINITIRNDGSTMEYVDGI 129  
Db 722 HYGNIKI-VNFILOHSAKVNAKTKNGYTAHQAOQGHTHIINVLLQNNAS-PNEL---- 775  
QY 130 FHLQNLNG-----VYQPS-----TAIYANTNSGTPTDLPNGVSWTTLVDP 174  
Db 776 ----TVNGNTALAIARRLGIVSVDTLKVYTBEMITT---TTEKHKRNVPETNMEVLD 828  
QY 175 MPD 177  
Db 829 MSD 831

RESULT 12  
US-11-096-568A-29105  
; Sequence 29105, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nikolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; PRIOR FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 29105  
; LENGTH: 946  
; TYPE: PRF  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)..(946)  
; OTHER INFORMATION: Ceres Seq. ID no. 4263811  
US-11-096-568A-29105

Query Match  
Best local Similarity 7.9%; Score 84.5; DB 7; Length 946;  
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;

QY 11 PNNASTVSATSLVNSTSSQVAKQEQNSSTSPAKSTNSLQHAQOAMTSSQSKRY 70  
Db 59 PNNSSVLSLSNPNLSSSFLPYVCNLTLESILVSNRKRSSIBEGFTNCERLIALKH 118  
QY 71 IPEHT-----YKVGDLKIRVNSLQ-QYKSVGYDQIGETANGAFWVINITI-RN 117  
Db 119 INFSTNKSTSPRGFGFSLAVLDPSSHVLSGNVGDYGFDPGLVQLRS-----LNLSPNLT 173  
QY 118 DGSTMEYVDGIFHLQNLNGVNYQDSTAETIYANTNSGTPTDLPNGVSWTTLVDFMDP 177  
Db 174 TGSVPVHLTKSLKLT-----EVSDNSLSTGTPGIXDYDELTL-LIDSLD 216  
QY 178 FWTYGVGHQVSLVASMG 195  
Db 217 NQJNG-----STPSLIG 228

RESULT 13  
US-11-096-568A-29104  
; Sequence 29104, Application US/11096568A  
; Publication No. US20060048240A1

GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: Thebydy  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 29104  
; LENGTH: 959  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(959)  
; OTHER INFORMATION: Ceres Seq. ID no. 4263810  
US-11-096-568A-29104

Query Match 7.9%; Score 84.5; DB 7; Length 959;  
Best Local Similarity 21.7%; Pred. No. 9.5;  
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;

QY 11 PRMASTVATTSVLNVTSSQVAKQONSSSTSPAKHSTNSLOHAQHOATTSSQSKLRY 70  
DB 72 PDNSSVISLSISNFDLSNSSFPLVVCNLTLESILDVSNRSLSPGCVTCERLIALKH 131  
QY 71 IPFHT-----YKVGDLERVNSLQ-QVKSVDGIGETANGAFVINITI-RN 117  
DB 132 LNFSTNKSTSPGFRGFSKLAVLDPSHNVLSGNVDYGFGLVQLRS-----LNFSTNRL 186  
QY 118 DGSTPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTIPTDLNPGVSMTNLVFDPMD 177  
DB 187 TGSVPVHLTKSLKLT-----EVSNDLSLSTIIPGICINDYQELT---LIDLSLD 229  
QY 178 FMTYGVQGHYSVLVSMG 195  
DB 230 NQLNG-----SIPSLIG 241

RESULT 14  
US-11-096-568A-29103  
; Sequence 29103, Application US/11096568A  
; Publication No. US20060048240A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexandrov, Nickolai et al.  
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides  
; TITLE OF INVENTION: Thebydy  
; FILE REFERENCE: 2750-1592PUS2  
; CURRENT APPLICATION NUMBER: US/11/096,568A  
; CURRENT FILING DATE: 2005-04-01  
; NUMBER OF SEQ ID NOS: 34471  
; SEQ ID NO 29103  
; LENGTH: 966  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(966)  
; OTHER INFORMATION: Ceres Seq. ID no. 4263809  
US-11-096-568A-29103

Query Match 7.9%; Score 84.5; DB 7; Length 966;  
Best Local Similarity 21.7%; Pred. No. 9.6;  
Matches 43; Conservative 32; Mismatches 82; Indels 41; Gaps 7;

QY 11 PRMASTVATTSVLNVTSSQVAKQONSSSTSPAKHSTNSLOHAQHOATTSSQSKLRY 70  
DB 79 PDNSSVISLSISNFDLSNSSFPLVVCNLTLESILDVSNRSLSPGCVTCERLIALKH 138  
QY 71 IPFHT-----YKVGDLERVNSLQ-QVKSVDGIGETANGAFVINITI-RN 117  
DB 139 LNFSTNKSTSPGFRGFSKLAVLDPSHNVLSGNVDYGFGLVQLRS-----LNFSTNRL 193

QY 118 DGSTPMEVVDGIFHLQNLGNVYQPDSTAEIYANTNSGTIPTDLNPGVSMTNLVFDPMD 177  
DB 194 TGSVPVHLTKSLKLT-----EVSNDLSLSTIIPGICINDYQELT---LIDLSLD 236  
QY 178 FMTYGVQGHYSVLVSMG 195  
DB 237 NQLNG-----SIPSLIG 248

RESULT 15  
US-10-485-517-146  
; Sequence 146, Application US/10485517  
; Publication No. US20050256299A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Biosynexis Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides  
; FILE REFERENCE: P100629MO  
; CURRENT APPLICATION NUMBER: US/10/485,517  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: GB 0118825.9  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: GB 0200349.9  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 424  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 146  
; LENGTH: 706  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-10-485-517-146

Query Match 7.8%; Score 84; DB 6; Length 706;  
Best Local Similarity 18.4%; Pred. No. 6.9;  
Matches 28; Conservative 35; Mismatches 67; Indels 22; Gaps 3;

QY 15 STVSATTSVLNVTSSQVAKQONSSSTSPAKHSTNSLOHAQHOATTSSQSKLRYIPFH 74  
DB 537 SAVSABETINQTTNPTMSVDDVNPATSAVTSNKAALCYEKLAOSKTDARAIALPHL 596  
QY 75 TYGKVGDLERVNSLQ-QVKSVDGIGETANGAFVINITIRNDG---STPMEVVDGIFH 131  
DB 597 NNAQADLVKSKINMASNLAGN-----TYKQGTDLNTMGNLQGAIAIN 639  
QY 132 LQ--NLNGNVYQPDSTAEIYANTNSGTIPTDL 161  
DB 640 DEQTLNQNQYODATPSSKTAVTVAVQAADI 671

Search completed: April 7, 2006, 16:40:51  
Job time : 15.5582 secs